

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2002, 11:29:28 ; Search time 66.93 seconds
(without alignments)
10470.527 Million cell updates/sec

Title: US-09-509-188-3

Sequence: 1 ggtacccacaaagaaacgcg.....ccgccttggttactt 2853

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/prodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/prodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/prodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/prodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/prodata/1/lna/CTUS.COMB.seq:*
6: /cgn2_6/prodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103.2	3.6	19124	2	US-08-487-826B-13
2	99.2	3.5	19124	2	US-08-487-826B-13
3	95	3.3	6124	4	US-08-213-419B-3
4	86.8	3.0	1511	1	US-07-991-867B-8
5	86.8	3.0	1511	1	US-08-107-755A-8
6	86.8	3.0	1511	2	US-08-544-332-8
7	86.6	3.0	7218	1	US-08-232-463-14
8	84.4	3.0	6124	4	US-08-213-419B-3
9	83.2	2.9	3440	5	US-08-471-791-27
10	83.2	2.9	3440	5	US-08-471-791-27
11	82.2	2.9	16442	2	US-08-883-891-208
12	80.6	2.8	665	2	US-08-883-891-208
13	79	2.8	9636	1	US-08-823-170B-1
14	79	2.8	9636	1	US-08-823-170B-1
15	77.6	2.7	615	4	US-08-954-441-1
16	76.6	2.7	3095	6	US-08-954-441-1
17	75.6	2.7	3095	6	US-08-954-441-1
18	75.4	2.6	1511	1	US-07-991-867B-8
19	75.4	2.6	1511	1	US-08-107-755A-8
20	75.4	2.6	1511	2	US-08-544-332-8
21	75.4	2.6	3489	2	US-08-728-323A-1
22	75.4	2.6	32207	2	US-08-770-379-20
23	75.4	2.6	32207	2	US-08-757-669A-20
24	75.4	2.6	32207	4	US-09-230-371A-20
25	75	2.6	837	4	US-08-998-416-288
26	75	2.6	51259	3	US-08-781-891-209
27	74.4	2.6	1431	4	US-09-316-083-2

28	74	2.6	732	4	US-08-998-416-1036	Sequence 1036, Ap
29	73.2	2.6	731	4	US-08-451-405A-2	Sequence 2, Appl
30	72	2.5	6152	4	US-08-973-463-1	Sequence 1, Appl
31	72	2.5	6768	1	US-08-107-755A-1	Sequence 1, Appl
32	72	2.5	8457	1	US-07-991-867B-1	Sequence 1, Appl
33	72	2.5	8457	1	US-08-544-332-1	Sequence 1, Appl
34	71.8	2.5	1241	1	US-07-593-657-6	Sequence 6, Appl
35	71.8	2.5	1241	4	US-08-942-012B-3	Sequence 3, Appl
36	71.6	2.5	636	4	US-08-998-416-1137	Sequence 1137, Ap
37	71	2.5	636	4	US-08-998-416-1137	Sequence 1137, Ap
38	71	2.5	854	4	US-08-998-416-534	Sequence 534, App
39	70.6	2.5	665	2	US-08-883-795A-36	Sequence 305, Appl
40	70.6	2.5	834	4	US-08-998-416-305	Sequence 786, App
41	70.4	2.5	711	4	US-08-998-416-786	Sequence 2, Appl
42	70.2	2.5	5852	1	US-07-867-106-2	Sequence 1, Appl
43	70.2	2.5	8920	2	US-08-446-855A-1	Sequence 1, Appl
44	70.2	2.5	8920	4	US-09-150-741-1	Sequence 1, Appl
45	70.2	2.5	12001	1	US-08-458-568A-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-08-487-826B-13
Sequence 13, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Slim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Knobe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13
Query Match 3.6%; Score 103.2; DB 2; Length 19124;

Best Local Similarity 42.68; Pred. No. 4,2e-10;
Matches 699; Conservative 0; Mismatches 926; Indels 15; Gaps 3;

```
OY 359 ttacttaatttngcngcatngtaccagtggttaacgmccttttttggancggt 418
Db 4 ttttttgcctatattatattatattatattatattatattatattatattat 63
OY 419 gattggttattgcttattggaatgngcncagtaagaatgaagaactattattg 478
Db 64 cttctatttttgcatttttttttttttccagttattattattattattttttgcatat 123
OY 479 gagaataataattataatgltcaatataagagaataataataatcactgagttacg 538
Db 124 tttttattattattattattattattattattattattattattattattattatt 183
OY 539 tatgagtagagtagaagatccttgaataatcttgagagactgaccttttccaaagt 598
Db 184 aattatccgagcagcaaaaaaattatattattattattattattattattattatt 243
OY 599 aaaaattgattatgtaacttaagtaacacatgaataataaaaaaataataatcaaa 658
Db 244 attatagtttccatttttttttttttttttttttttttttttttttttttttttt 303
OY 659 tagaanaaacgtatgtagtacccttcaagctttgacttattcttggttcaacccc 718
Db 304 ttaattatattattattattattattattattattattattattattattattatt 355
OY 719 taaaccttaagttccaaacataaaattcattatgcatcttcttcttcttcttga 778
Db 356 taaatttttttttttttttttttttttttttttttttttttttttttttttttttt 415
OY 779 aagtgaaacaaatattatcaagttattatgcttttccaaataaagaataaaaaaa 838
Db 416 aaaaaaaaaaaaaattattattattattattattattattattattattattattatt 475
OY 839 acaaaaataatagtagtacaataaaaaaataaattattattattattattattatt 898
Db 476 ttaattattattattattattattattattattattattattattattattattatt 535
OY 899 actaaaccttaaaccttaataataaccttgaagtaaaccttgaagtaaaccttga 958
Db 536 tgaagaagactttgagatattattattattattattattattattattattattatt 595
OY 959 ttaaacattaaacattaaacataaaccttaaaccttaaaccttaaaccttaaacctt 1018
Db 596 taaattattattattattattattattattattattattattattattattattatt 655
OY 1019 taaagtttagtttttttttttttttttttttttttttttttttttttttttttt 655
Db 656 tcaacatggaatatttttttttttttttttttttttttttttttttttttttttttt 715
OY 1079 aagagttatggttaggattagacttaggatttaggtttaggtttaggtttaggtt 1138
Db 716 acttattattattattattattattattattattattattattattattattattatt 775
OY 1139 gtaatttttttttttttttttttttttttttttttttttttttttttttttttttt 1198
Db 776 ctgctatttttttttttttttttttttttttttttttttttttttttttttttttt 832
OY 1199 ttataataaataataataataataataataataataataataataataataataata 1258
Db 833 attcaattattattattattattattattattattattattattattattattattatt 892
OY 1259 aacattatttttttttttttttttttttttttttttttttttttttttttttttttt 1318
Db 893 gcaacattattattattattattattattattattattattattattattattattatt 952
OY 1319 atttttcagctaaataataataataataataataataataataataataataataata 1378
Db 953 atttcaatttttttttttttttttttttttttttttttttttttttttttttttttttt 1012
OY 1379 aaatgctatattggtcaaacatggttaataataataataataataataataataata 1458
```

```
Db 1013 ttatattattattattattattattattattattattattattattattattattatt 1068
OY 1439 ttcttaactggttgaataaagtcgaagtgaaatlltagaagcaatgagtaatat 1498
Db 1069 tatcatattattattattattattattattattattattattattattattattattatt 1128
OY 1499 ttgttcaatttttttaagaatcagatattgataatacaaatattattattattattatt 1558
Db 1129 attattattattattattattattattattattattattattattattattattattatt 1188
OY 1559 cctaaataatccttaagaagtgaaagcgaagaataagtcacggttttgggaaagccta 1618
Db 1189 aaaaattctattattattattattattattattattattattattattattattattattatt 1248
OY 1619 acttggcccaagtcacaaatccttctgattattatcaaatccttaacatttagt 1678
Db 1249 attattattattattattattattattattattattattattattattattattattatt 1308
OY 1679 tagagtttaagacacacacacacacacacacacacacacacacacacacacacacac 1738
Db 1309 aagacttaagcaaaacacacacacacacacacacacacacacacacacacacacacac 1368
OY 1739 gcgaaataataaatcttcttattattattattattattattattattattattattattatt 1798
Db 1369 atttattattattattattattattattattattattattattattattattattattatt 1428
OY 1799 actaaacaaatttaataaagaataaaggacacacacacacacacacacacacacacacac 1858
Db 1429 tcaattattattattattattattattattattattattattattattattattattattatt 1488
OY 1859 ttcttaagatggttgaataataataataataataataataataataataataataataata 1918
Db 1489 ttttttccattattattattattattattattattattattattattattattattattattatt 1548
OY 1919 tatcttgcagtcgataatcattacattacattacattacattacattacattacattacatt 1978
Db 1549 aacatttcaaaaaaatttgaacacacacacacacacacacacacacacacacacacacacac 1608
OY 1979 ttattattattattattattattattattattattattattattattattattattattatt 1628
Db 1609 ttcttattattattattattattattattattattattattattattattattattattatt
```

RESULT 2
US-08-487-826B-13/C
Sequence 13, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chien, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-Zhaun
APPLICANT: Williams, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobe Martens Olson & Bear
STREET: 620 Newport Center Drive
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435

QY 374 gngcangtcatcaggtttagtgcnccttttttttgancggttgatgttgcctt 433
 Db 747 taaactatattatgttaattcttttttttttttttttttttttttttttttt 1111
 QY 434 attgagatgtgcmagcgaagaaatgaaagacatttatattggg-----gaaatat 488
 Db 807 attatatacatattatgaaanaataattccatttaggaatggaaggtcttcatg 866
 QY 489 aatttaattgttcaatatatagaagaaatattatncccttgatgcttgcgtatgagcg 548
 Db 867 gatttaaaaaaataatataatataatataatataatataatataatgagtgatgaa 926
 QY 549 agtgaagatccttggaaatatttggaaacttgcctttcttcaaaaaagtaaatattg 608
 Db 927 acatatataatttttttaatatatttgaagagaaagaaatgaaatgcatattag 986
 QY 609 atagttaacttaagttacacatgaaatgaaanaaaatgaaatgaaatgaaanaac 668
 Db 987 taatagctatataactaataatgttatttcatccttttaatttagctatataatata 1046
 QY 669 tgaatgagctacaccctcaagcttggacttctgttgcctccctcaacctcta 728
 Db 1047 atacatatataatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 1106
 QY 729 agtgcac 788
 Db 1107 ttaagaaccttataaataatgttataaataatgttataaanaaaatataatattggt 1166
 QY 789 aaatattacagttatataatgttttttcaaaaaagataaaaaaataaaaaata 848
 Db 1167 aagtaactgagaaatattttttccatcatattttttccatataatgaaatgaaatgaa 1226
 QY 849 atagatgttacaac 906
 Db 1227 ggtatacttataatataatataatgaaatgaaatgaaatgaaatgaaatgaaatg 1286
 QY 907 ctaaacaccttaataacttttaggtgaaaccccaacacacacacacacacacacac 966
 Db 1287 ataaaaatagtaagacttttagtgagaaagacacacacacacacacacacacacac 1346
 QY 967 taaac 1026
 Db 1347 ttgttaataataaataaataaataaataaataaataaataaataaataaataaata 1406
 QY 1027 tagtgtttgattatagtttaggttttccaaagtttaagtttaccacacacacac 1086
 Db 1407 aatgaatttcttctgttgaatattatattatgatatgatttctcgagaccacacacac 1466
 QY 1087 atgtttgaggtatgacttaggttaggttaggttaggttaggttaggttaggttaggt 1146
 Db 1467 aaaaatcagcttagtctgataatcaaatgatatataataataatatttttttttaac 1526
 QY 1147 taaaaaataatttttttgaacacacacacacacacacacacacacacacacacacac 1199
 Db 1527 tccctagaatattttaaagaataataatcgacacacacacacacacacacacacacac 1586
 QY 1200 ttataataaataaataaataaataaataaataaataaataaataaataaataaata 1259
 Db 1587 ttaatttggaagaaaaaataatttatataattattattgttaatttatatatata 1646
 QY 1260 acattattttttgttcaaaaaaattgtcactttgaattcccaatgcaaaaaattatta 1319
 Db 1647 ttataattaaatt 1706
 QY 1320 tttttcagcnaaatgaatgcaagtgatgacttcaataaataatttttttttttttttt 1379
 Db 1707 tgaatcacaatattttaaataatttttttttttttttttttttttttttttttttttt 1766
 QY 1380 aatgcataattgttcaac 1439
 Db 1767 ataactaaccttgttattctttaaanaaaatgtacacacacacacacacacacacacac 1826

QY 1440 ttcttaactctgtgta-----aaatgtcaagtaaaattatttttagaagcaatttgatg 1494
 Db 1827 atataatataatgtatttcttcaatttaataatttttagatatttatatttttttttt 1886
 QY 1495 atatttgcattttttttaaagatctgagatgtgaataaacaacattttatgtatg 1554
 Db 1887 gaataacataattttgataac 1946
 QY 1555 tgaac 1614
 Db 1947 ctccatttatataatataatataatataatataatataatataatataatataatg 2006
 QY 1615 gctaacatagcccaagatcaaaatcttcttttttttttttttttttttttttttttttt 1674
 Db 2007 ttttaataatagtt 2066
 QY 1675 tagttagagttaaatgac 1734
 Db 2067 attttttttgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 2126
 QY 1735 actagcaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 1794
 Db 2127 aatgtac 2184
 QY 1795 atatacnaaac 1854
 Db 2185 atataatataatataatataatataatataatataatataatataatataatataat 2244
 QY 1855 gttattcttcaagatgagatgagatgagatgagatgagatgagatgagatgagatgagatg 1914
 Db 2245 aaatataatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagatg 2296
 QY 1915 atataatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagatg 1974
 Db 2297 atattataaatacattataatataatataatataatataatataatataatataatata 2356
 QY 1975 atatttttatataatataatataatataatataatataatataatataatataatataat 2034
 Db 2357 aaaaatgaatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagatg 2416
 QY 2035 acttttcttctgttgcac 2092
 Db 2417 atattccttgttttcttattgttgaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 2474

RESULT 4
 US-07-991-867B-8
 Sequence 8, Application US/07991867B
 Patent No. 5476781
 GENERAL INFORMATION:
 APPLICANT: Moyer, Richard W.
 APPLICANT: Hall, Richard L.
 APPLICANT: Gruidl, Michael E.
 TITLE OF INVENTION: No. 5476781e1 Entomopoxvirus Expression System
 NUMBER OF SEQUENCES: 66
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: David R. Saliwanichik
 STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: FL
 COUNTRY: USA
 ZIP: 32606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/991,867B
 FILING DATE: 12-DEC-1992
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO 92/14818

FILED DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanichik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UFI114.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Amsacta moorei entemopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS
LOCATION: 852..1511
US-07-991-867B-8

Query Match 3.0%; Score 86.8; DB 1; Length 1511;
Best Local Similarity 43.8%; Pred. No. 2.2e-07;
Matches 476; Conservative 0; Mismatches 607; Indels 5; Gaps 2;

QY 915 taaatttaacttttagttaaacccttaaccccttgataaacttaacataaacatt 974
DB 196 TAATTAATTAATCACTTATTTGCAATTAATGAATTAATAATTAATTAATCAAT 255
QY 975 aaacacctaaccctaactcctaactcctaaccctaagtglttaagtgtagttt 1034
DB 256 AATCTACACAGAGAACATTAATCTTGCCACTATTTCATTTTGCATTTTATTA 315
QY 1035 ttgattataagtttaggattatccaaaggttaaggtttacccaaggtttagttta 1094
DB 316 TGTATTTTAAATGTAAGAGCATCTTATTAACAAATTAATGACATTAAGCTTAATTT 375
QY 1095 gggattatc---gacttagagattagtggttactgcagcagcttcaagatttttaaaa 1151
DB 376 TTTATTTTCTTACCTTAGAATTAATTTTGATATGAATTAATTAATTTCTGTTAAAG 435
QY 1152 aatatttttttgtaaacactactattttatttttttttttttttttttttttttttttt 1211
DB 436 TCACATTTTAATCCAGCAACAATACTTTTATTTTATTAATGACATTTTATACAAAAT 495
QY 1212 cataatataatthaactcactcgtttcactatataagttgtaagttgtaacattttttt 1271
DB 496 TGTCTAAATCACTTTCTCAAAAATTAAGCACTCATATGACCAATTAATATCATATTA 555
QY 1272 tgttaacaaaaaattcactcacttagaattcacaatgcaaaaattatttttttcagctaa 1331
DB 556 TCTACATATGATGATTCATTAATTAATTAATTTGTTTAAATGAATTAATAATTTTATTA 615
QY 1332 aatttaagtgcaagtgcatgactcattataaataattttattttatcccaaatgctatattg 1391
DB 616 AATATATTTCCGATCATGATTTATTAATTTTATTAATAATCTATTAATCTAATTAATGA 675
QY 1392 gtcaaacatgtgtaatttaataagaacttaattatcatttttttttttttttttttttt 1451

DB 676 GTTATTAATTACACTTTTTCATTAATAAATATATCTAATTAATTTTCGACATTCCT 735
QY 1452 gtaaaatgtcaaaagtaaaattatttagaaacgaattgagtaaatgttttgcatttt 1511
DB 736 GTTGTTCCTCCGAAAAACATAGGACCAATTTAATTTATGACATTTTATTAATTA 795
QY 1512 ttaaaaga--tatcgatatgaataacacaaattttattgtagtgaagcaacaaatttc 1569
DB 796 TTGATATATTTTTCAAAAAATTAATCAATGAAAAAATAAATTAATCAAAAATG 855
QY 1570 atcctaagaagtgaaacgaataagtcacglttttggggaagctaatatgcccac 1629
DB 856 ATTTACTAAATCTGATTAATTAATTTAATTAATTTAATTAATTTAATTTAATTTA 915
QY 1630 aagtcataaaactttcttgattatccaaatcccttaacaaatttagttaggttaata 1689
DB 916 TAATTAATTAACAGATTAATGTTAATTAATTAATTAATTAATTAATTAATTTAG 975
QY 1690 gaccacaacatgattatcatcatatagatatcttaaaaaacttaactgagcaataata 1749
DB 976 AACAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1809
QY 1750 aaatcttcttattatcacaatcccttaataaaacttaattatataataaacaat 1869
DB 1036 AAGTTTATATATTTCAAAATTTAATTAATTAATTAATTTAATTAATTAATTAATTA 1899
QY 1810 ttaatttaaaagaaataagggacatgatacataataaataatatttttttaagat 1929
DB 1096 ATATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1975
QY 1870 agtgaataattatataaccagtcacatatatttatacaaaataaataattttcgta 1929
DB 1156 CACATTTCTATAGAAATTTTAAATTTGCAATTCATTAATTAATTAATTAATTAATTA 1975
QY 1930 gtccgataatcattactataaattcaaaacacatgtagatgattattttatatat 1989
DB 1216 ATATTTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1989
QY 1990 atatatat 1997
DB 1276 ATGTTTTT 1283

RESULT 5
US-08-107-755A-8
Sequence 8, Application US/08107755A
Patent No. 5721352
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5721352e1 Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanichik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/107,755A
FILING DATE: 19-AUG-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,658
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/657,584
 FILING DATE: 19-FEB-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Saliwanchik, David R.
 REGISTRATION NUMBER: 31,794
 REFERENCE/DOCKET NUMBER: UFI14.C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (904) 375-8100
 TELEFAX: (904) 372-5800
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1511 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: Amsacta moorei entomopoxvirus
 FEATURE:
 NAME/KEY: CDS
 LOCATION: complement (18..218)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: complement (234..782)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 852..1511
 US-08-107-755A-8

Query Match 3.0%; Score 86.8; DB 1; Length 1511;
 Best Local Similarity 43.8%; Pred. No. 2.2e-07;
 Matches 476; Conservative 0; Mismatches 607; Indels 5; Gaps 2;

QY 915 taaatgaacttttagtgaacccttaacccttgatgaatccttaacataacatt 974
 DB 196 taaatgaacttttagtgaacccttaacccttgatgaatccttaacataacatt 974
 QY 975 aaaaactaaacccttaacccttaacccttgatgaatccttaacataacatt 1034
 DB 256 aaaaactaaacccttaacccttaacccttgatgaatccttaacataacatt 1034
 QY 1035 ttgatttagttagtgaacccttaacccttgatgaatccttaacataacatt 1094
 DB 316 ttgatttagttagtgaacccttaacccttgatgaatccttaacataacatt 1094
 QY 1095 gggattatcctgaactgaactgaactgaactgaactgaactgaactgaact 1151
 DB 376 ttgatttagttagtgaacccttaacccttgatgaatccttaacataacatt 1151
 QY 1152 aataatcttttgaactgaactgaactgaactgaactgaactgaactgaact 1211
 DB 436 tcaacatttttgaactgaactgaactgaactgaactgaactgaactgaact 1211
 QY 1212 cataataaactgaactgaactgaactgaactgaactgaactgaactgaact 1271
 DB 496 tggcttaaacatttttgaactgaactgaactgaactgaactgaactgaact 1271
 QY 1272 tggtaaaaaaactgaactgaactgaactgaactgaactgaactgaactgaact 1331
 DB 556 tggtaaaaaaactgaactgaactgaactgaactgaactgaactgaactgaact 1331
 QY 1332 aataatgaactgaactgaactgaactgaactgaactgaactgaactgaact 1391
 DB 616 aataatgaactgaactgaactgaactgaactgaactgaactgaactgaact 1391
 QY 1392 gtaacacatgtaactgaactgaactgaactgaactgaactgaactgaactgaact 1451
 DB 676 gtaacacatgtaactgaactgaactgaactgaactgaactgaactgaactgaact 1451
 QY 1452 gtaacacatgtaactgaactgaactgaactgaactgaactgaactgaactgaact 1511
 DB 736 gtaacacatgtaactgaactgaactgaactgaactgaactgaactgaactgaact 1511

QY 1512 ttaaga--tatgaatataaataaacaatttttgatgatgaaccataaattc 1569
 DB 796 ttgatttatattttttcaaaaaaaatttatcattgataaaaaaaatttatcatt 855
 QY 1570 atcctaagaagtgtaaccgaataaagtgtaaccgttttggaaggaactatgcacca 1629
 DB 856 atttactaaatttgcatttatatttttatatttatatttttatatttttatatt 915
 QY 1630 aagtcataaatttcttcttgatatttcaaaatccttaacaaattagtgatgaata 1689
 DB 916 taaatgaacacagatgaactgaactgaactgaactgaactgaactgaactgaact 975
 QY 1690 gaccaacacatgatactcatatagatgaactgaactgaactgaactgaactgaact 1749
 DB 976 aagaattgcatatatttatatttatatttatatttatatttatatttatatt 1035
 QY 1750 aaatcttcttatttataaataaactgaactgaactgaactgaactgaactgaact 1809
 DB 1036 aaatcttcttatttataaataaactgaactgaactgaactgaactgaactgaact 1095
 QY 1810 ttaataaagaagaataagggacatgatacataaataatagttatttcttaagat 1869
 DB 1096 ataatgaacttttagatatttttatatttatatttatatttatatttatatt 1155
 QY 1870 agtgataatataatatacagagcagcagcagcagcagcagcagcagcagcagc 1929
 DB 1156 cacattctatagaaatttttaatttgatgaactgaactgaactgaactgaact 1215
 QY 1930 gtccgataactgaactgaactgaactgaactgaactgaactgaactgaactgaact 1275
 DB 1216 ataatgttatgaatttttaaaaaaaatttatatttatatttatatttatatt 1275
 QY 1990 atatacat 1997
 DB 1276 atgttttt 1283

RESULT 6
 US-08-544-332-8
 Sequence 8, Application US/08544332
 Patent No. 5935777
 GENERAL INFORMATION:
 APPLICANT: Moyer, Richard W.
 APPLICANT: Hall, Richard L.
 APPLICANT: Gruidl, Michael E.
 TITLE OF INVENTION: No. 5935777e1 Entomopoxvirus Expression System
 NUMBER OF SEQUENCES: 77
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Gerard H. Bencen
 STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: FL
 COUNTRY: USA
 ZIP: 32606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/544,332
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/991,867
 FILING DATE: 07-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/107,755
 FILING DATE: 19-AUG-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO 92/14818
 FILING DATE: 12-FEB-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bengen, Gerard H.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UFI14.C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Amacta moorei entemopovirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS
LOCATION: 852..1511
US-08-544-332-8

Query Match 3.0%; Score 86.8; DB 2; Length 1511;
Best Local Similarity 43.8%; Pzed. No. 2.2e-07;
Matches 476; Conservative 0; Mismatches 607; Indels 5; Gaps 2;

QY 915 taaatttaaaccttttagttaaacccttaacccttgatgaataacctaataacatt 974
DB 196 TAAATACAT 255
QY 975 aaaaacacaaacctaataacctaataacctaataacctaataacctaataaccta 1034
DB 256 AATCTACACACAGACATATATATATATATATATATATATATATATATATATAT 315
QY 1035 ttagttaaagtttagttaaagtttagttaaagtttagttaaagtttagttaaagtt 1094
DB 316 TGTATTTTAT 375
QY 1095 ggaattatc--gacttagagtttagttaaagtttagttaaagtttagttaaagtt 1151
DB 376 TTTATTTTCTTCTTATAGATATATATATATATATATATATATATATATATATAT 435
QY 1152 aatattttttttagttaaactactattttatttttttttttttttttttttttt 1211
DB 436 TCACAATTTAATCCACCAATATATATATATATATATATATATATATATATATAT 495
QY 1212 cataataatttaataacccatctgtttcatatttaagtttagttaaagtttagt 1271
DB 496 TGTCTTAAATCATTTCTTCAAAAAATGACACTCATATATATATATATATATATAT 555
QY 1272 tgttaaaaaaaatctgcacttagaattccaaatgcaaaattattatttttcagctaa 1331
DB 556 TCTACATATATGATTTATATATATATATATATATATATATATATATATATATAT 615
QY 1332 aatttaattgcaaatgcatgactataataattattatttttttttttttttttttt 1391
DB 616 AATATATTTCCGTCATGATTTATATATATATATATATATATATATATATATATAT 675
QY 1392 gtcaaacatgctgaataataataataattattatttttttttttttttttttttt 1451
DB 676 GTTATATTTACACATTTTGTATGATATAAATATATATATATATATATATATATAT 735

QY 1452 gtataaatgtcaaatgaataattattagaacgaattgaataatttttttttttttt 1511
DB 736 GTTGTCTTGGCAGAAAAACATAGACCAATATATATATATATATATATATATATAT 795
QY 1512 ttaaaaga--talcgaatatgaataacaaattttatttttttttttttttttttttt 1569
DB 796 TTGATATATTTTTCACAAAAAAATTAATGAAAGAAAAAAATTAATATATATATAT 855
QY 1570 atcctaagaagtgcgaacgcaagaataagttcaagttttgggggaactacttgccca 1629
DB 856 ATTTACTAAATCTGAT 915
QY 1630 aagtcatacaaatcttctgtattatcaaaatcccttaacaaattagtttagttaaata 1689
DB 916 TATATATTAACAGAT 975
QY 1690 gaccaaacaatgattatcatatagaaataatctcaaaaataactagcgaataatta 1749
DB 976 AAGCAATTCAT 1035
QY 1750 aaatcttcttattattatcaaaatccctataaaaactattattattattattatt 1809
DB 1036 AAGTTTAT 1095
QY 1810 ttttaataaagaataaaggaacatgatacaataaataatgatttttttttttttttt 1869
DB 1096 ATATACATATTTAGAT 1155
QY 1870 agtgaataattatataataccagtcacataattttcaaaaataataataatttcgta 1929
DB 1156 CACATTCAT 1215
QY 1930 gtccgataatcattactataataattcataaacacatgtagattatttttttttt 1989
DB 1216 ATATATTTAGTAAATTTAAAAAATTAATATATATATATATATATATATATAT 1275
QY 1990 atatacat 1997
DB 1276 ATGTTTTT 1283

RESULT 7
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991

NAME/KEY: CDS

[illegible][illegible]

QY 1229 ctccatcgtttcaatgaagtcgctgtgaacattatttttgcacaanaatgct 1288
Db 1665 AAAAA-----ATTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 16133
QY 1289 cacttgaattcccaatgcaaaatttatttatttcttcagctcaaaatgaatgcaatgctc 1348
Db 1612 TAAATTTTTTTTTTCTCCAAATATAAAAGTTTTTCATACAGATTAATATAGTCTG 15533
QY 1349 atgacttataataatttatttatttctccaatgctatattgtgccaacatgtaatt 1408
Db 1552 ATTTATTTTCATTAAATATTTTCAGGAGTTAAAAAAATTAATTTATATATCCAT-ATG 14943
QY 1409 aatgaacctaatattatttatttatttcttatttcttcaatctgtgcaaaatgcaaatga 1468
Db 1493 AATTATTCAGACATTAAGCTGATTTTTTTTTTTCTTGCTGACAGAAAATCATATACATA 14344
QY 1469 aaattattagaacgaattgtaataatttgtttcattttttaaagaatcgaata 1528
Db 1433 ATATATTTTACACAAAAGAA-----AATTCATTAATCTTAATTTACTAGACCTTATATTA 1380
QY 1529 tgaataaccaatttatttatttatttatttgccttaaaatcttctaagaagtgaaagc 1568
Db 1379 TTTTATTTTTTTTTTATTTATTTCTTAATTAACAATAATATAAGTGCATGATTAATATTT 1320
QY 1589 aagataagcaacgcttttgggaagctcaactatgcccgaagccatcaaatcttct 1648
Db 1319 GCTTTTCCATCATTAAGCTCTTACTAATTTTATGTTTCATGTAATATTTTAATAATATCA 1260
QY 1649 tgaattatcaaaatcccttacaatttgaatttagttagtlaagacacaacatgattc 1708
Db 1199 ATATGATGGAATAAAATTTTATCTTAGTACTTCCATTATTAATTAATTTTTTTTTTAATA 1140
QY 1709 atcattatgaatatctcaaaaataattcaggaataataaactcttcttattat 1768
Db 1259 TTTATCTTAATTAATGATTTTATGAGTAATACCTTTTTTTCTTATTAATGAATAAT 1200
QY 1769 caaaatccctataaaactatttataat-atactaaacaatttcaatuaaagaanaa 1827
Db 1139 GCATTTTAATACATTTTTCATATAGTCTCTTAATAAATAATTTGTGGTAATACACCC 1080
QY 1828 agggacacatgatacaataaataatgtc-tattcttaagaatgataatataat 1866
Db 1079 -AATCATTTTAAATTAATTTCTATATATATGTAATATATATATATAGACGTAAATTAATAAG 1020
QY 1887 ataccagtcataattattaca-----ataaataatttlogtgcgataaca 1941
Db 1019 ATGAAAAATACATATTAGTTAGTAATAGCAATTACATATATGACATTTTCAATTTCT 960
QY 1942 ttactataaattcaataaacacatgtaagtataatttattatataatataa 2001
Db 959 CTTTAAATATATTTTAAATAAATAATATATATATGTTTCATTAACGTCAATAATATATATAT 900
QY 2002 a 2002
Db 899 A 899

RESULT 9
US-08-471-791-27/C
Sequence 27, Application US/08471791
Patent No. 5723595
GENERAL INFORMATION:
APPLICANT: Thompson, Gregory A
APPLICANT: Knafit, Vlc C
TITLE OF INVENTION: Plant Desaturases-Compositions
TITLE OF INVENTION: and Uses
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California

COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoftword 5.1 (a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,791
FILING DATE: 6-JUNE-95
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/762,762
FILING DATE: 16-SEPT-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/01746
FILING DATE: 14-MAR-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/615,784
FILING DATE: 14-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/567,373
FILING DATE: 13-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/494,106
FILING DATE: 16-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Lassen, Elizabeth
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 69-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
TELEX: 350370 CGNE
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 3440 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
US-08-471-791-27

	Query Match	2.9%	Score 83.2:	DB 1:	Length 3440:
	Best Local Similarity	52.2%:	Pred. No. 1.1e-06:		
	Matches 239:	Conservative 0:	Mismatches 208:	Indels 11:	Gaps 2:
OY	1162 tgacttagattagtggttttctacgacgcgtccaaagtatttttaaaaatatattttt	1161			
Dd	457 TAAATTTTTTAAAAAGCTGTGCTTAATTGGTGTTGATTATTAAATTCATATAAATCTT	398			
OY	1162 ttgtacaacacctatcttttattatattttttacccttttatatacaaacataataa	1221			
Dd	397 TTGTATTGTCACACAATTTTCTAACCAATTAATTTTTTTATTATTATTAATAAAGCTGA	338			
OY	1222 ttaataactcactcgtttcatatataagtygaatctgtacatattttttgttcaaaa	1281			
Dd	337 GATATTTCTGTTTTTTTTTTTAAT-----CTGTGTACAAAATATTGTTACACAA	285			
OY	1262 aaattgctactttagaatccaatgccaaaattattatttttcsgctcaaaattaatgc	1341			
Dd	284 AAAAGTCACTTAAATTAATTTCAATACAAATTAATTAATTTTCACTGGAATTAATGT	225			
OY	1342 aaagcgctgtagctcataaataattattattatctatcaaatgctctatattgcccacaacg	1401			
Dd	224 AAACGCATATGATTTTAAACATACAGAGACCCTGACGCCGCCACGCGGATATGAANT	165			

COUNTRY: USA
 ZIP: 98104-7092
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/781,891
 FILING DATE: 27-DEC-1996
 CLASSIFICATION: . 800
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6090620tenbmg Ph.D. Carol
 REGISTRATION NUMBER: 39,317
 REFERENCE/DOCKET NUMBER: 240052.419
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 208:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 16442 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

Query Match 2.9%; Score 82.2; DB 3; Length 16442;
Best Local Similarity 60.5%; Pred. No. 2.2e-06;
Matches 135; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 2514 tctactcaactctctctctcaagagctcttaacatgctctctctcaagagctctta 2573
DB 16192 tct 16251
QY 2574 catgctctctcaagctctctcaagagctctcttaagctctcttaagctctctta 2633
DB 16252 cct 16311
QY 2634 catgctctctcaagctctctcaagagctctctcaagagctctctcaagagctctta 2693
DB 16312 cct 16371
QY 2694 catgctctctcaagctctctcaagagctctctcaagagctctctcaagagctc 2736
DB 16372 cct 16414

RESULT 12

US-08-883-795A-36/c
Sequence 36, Application US/08883795A

Patent No. 5985607

GENERAL INFORMATION:

APPLICANT: Delcuve, Genevieve

TITLE OF INVENTION: Recombinant DNA Molecules and Expression

NUMBER OF INVENTION: Vectors for Tissue Plasminogen Activator

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: BERESKIN & PARR

STREET: 40 King Street West

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5H 3Y2

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/883,795A

FILING DATE: 27-JUN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Gravelle, Micheline

REGISTRATION NUMBER: 40,261

REFERENCE/DOCKET NUMBER: 7841-062

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 364-7311

TELEFAX: (416) 361-1398

INFORMATION FOR SEQ. ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 665 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORGANISM: Homo sapiens

IMMEDIATE SOURCE:

CLONE: Rh 32

US-08-883-795A-36

Query Match 2.8%; Score 80.6; DB 2; Length 665;
Best Local Similarity 46.8%; Pred. No. 2.4e-06;
Matches 286; Conservative 0; Mismatches 324; Indels 1; Gaps 1;

QY 1215 aatataatttaactccactcgtcttcaataaagtgatcattgtaacattatttctgt 1274
DB 626 AATACAAATTCGAGATCAATACACCGCTTTTAAATTTTCCGCTTTTAACTA 567

QY 1275 taacaaaataatgtaacattgtaacattgtaacattgtaacattgtaacattgtaac 1334
DB 566 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 507

QY 1335 taatgtaacattgtaacattgtaacattgtaacattgtaacattgtaacattgtaac 1394
DB 506 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 447

QY 1395 aaacatgtaacattgtaacattgtaacattgtaacattgtaacattgtaacattgtaac 1454
DB 446 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 387

QY 1455 aaatgtaacattgtaacattgtaacattgtaacattgtaacattgtaacattgtaac 1514
DB 386 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 327

QY 1515 aaagatcgaatagaaatgtaacattgtaacattgtaacattgtaacattgtaacattgtaac 1574
DB 326 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 268

QY 1575 aaagatcgaatagaaatgtaacattgtaacattgtaacattgtaacattgtaacattgtaac 1634
DB 267 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 208

QY 1635 atcaaaatctctgtaacattgtaacattgtaacattgtaacattgtaacattgtaacattgtaac 1694
DB 207 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 148

QY 1695 aaacatgtaacattgtaacattgtaacattgtaacattgtaacattgtaacattgtaacattgtaac 1754
DB 147 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 88

QY 1755 ttctttattatcaaaatccttaaaacttaaaacttaaaacttaaaacttaaaacttaaaacttaaaact 1814
DB 87 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 28

QY 1815 ttaaaagaaa 1825
DB 27 AATAAAACACA 17

RESULT 13

US-08-323-170B-1/c
Sequence 1, Application US/08323170B

Patent No. 5733772

GENERAL INFORMATION:

APPLICANT: Williamson, Kim C.

TITLE OF INVENTION: Cloning and Expression of Plasmidum

NUMBER OF INVENTION: falicparum Transmission-Blocking Target Antigen, Pf230

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/323,170B

FILING DATE: 13-OCT-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/010,409

1 FILING DATE: 29-JAN-1993
 2 ATTORNEY/AGENT INFORMATION:
 3 NAME: Einhorn, Gregory P.
 4 REGISTRATION NUMBER: 38,440
 5 REFERENCE/DOCKET NUMBER: 015280-11311005
 6 TELECOMMUNICATION INFORMATION:
 7 TELEPHONE: (415) 576-0200
 8 TELEFAX: (415) 576-0300
 9 INFORMATION FOR SEQ ID NO: 1:
 10 SEQUENCE CHARACTERISTICS:
 11 LENGTH: 9636 base pairs
 12 TYPE: nucleic acid
 13 STRANDEDNESS: single
 14 TOPOLOGY: linear
 15 MOLECULE TYPE: DNA (genomic)
 16 FEATURE:
 17 NAME/KEY: CDS
 18 LOCATION: 149..9556
 19 15-08-05

Query Match	2.88;	Score 79;	DB 4;	Length 9636;
Best Local Similarity	59.6%;	Pred. No. 7.4e-06;		
Matches 133;	Conservative 0;	Mismatches 90;	Indels 0;	Gaps 0

Accession	Sequence	Position
QY 2507	tgaggagtaacttaactcctcctctcaagaagctccttaatagtccctcttcaagc	2566
Db 1492	TTTTTCACTTCACTACCTTTCTTTCTTCACTACCTTTCACTTCTTCACTAC	1433
QY 2567	tccttacaatgcctcttacaatgcctctcaagaagctcctttaatgcttccttaagc	2626
Db 1432	CTCTTCACTACCTTTCACTACCTTTCTTTCACTACCTTTCACTACCTTTCACTAC	1373
QY 2627	tccttacaatgcctcttacaatgcctctcaagaagctcctttaatgcttccttaagc	2686
Db 1372	CTCTTCACTACCTTTCACTACCTTTCTTTCACTACCTTTCACTACCTTTCACTAC	1313
QY 2687	cccttacaatgcctcttacaatgcctctcaagaagctcctttaatgcttccttaagc	2729
Db 1312	CTCTTCACTACCTTTCACTACCTTTCTTTCACTACCTTTCACTACCTTTCACTAC	1270

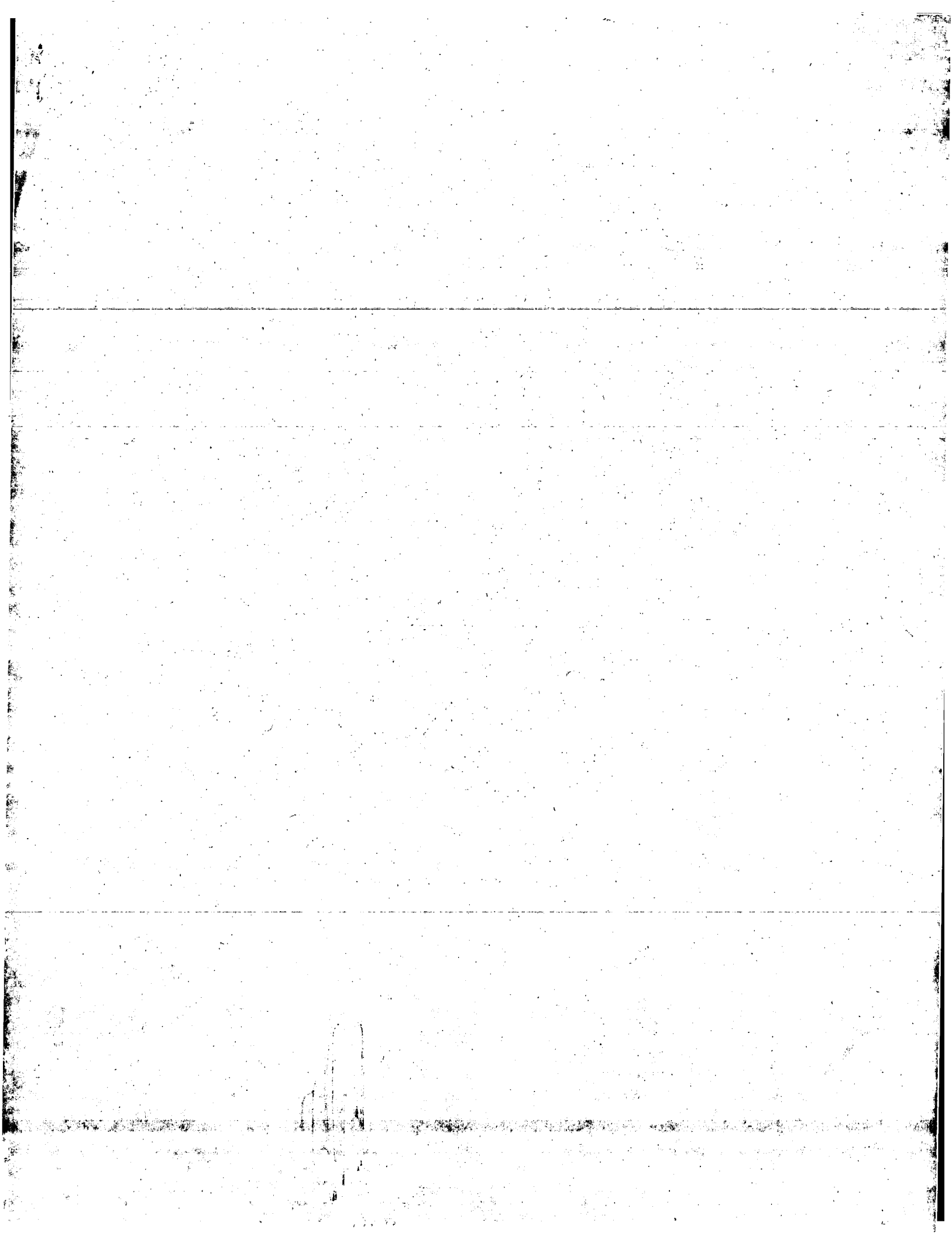
RESULT 15
US-08-998-416-186/c
Sequence 186, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:

APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jürgen
APPLICANT: Knechtle, Philipp
APPLICANT: Rebschund, Corinne

TITLE OF INVENTION: GENOMIC DNA SEQUENCES
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30C
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/998,416
 FILING DATE: 24-DEC-1997
 CLASSIFICATION: 435



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2002, 11:26:23 ; Search time 3471.84 Seconds

(without alignments)
17196.475 Million cell updates/sec

Title: US-09-509-188-3
Perfect score: 2853
Sequence: 1 ggtaccaccaagaacacg.....ccgtcttgatttaactt 2853

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBankl:*

- 1: gb_da:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_cm:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_da:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	2844	99.7	2853	6	A97606	A97606 Sequence 3
2	2840.8	99.6	2966	8	AF136223	AF136223 Brassica
3	462	16.2	674	6	A97605	A97605 Sequence 2
4	393.2	13.8	497	6	A97604	A97604 Sequence 1
5	184.4	6.5	75188	8	AB005238	AB005238 Arabidops
6	175.8	6.2	2966	8	AF136223	AF136223 Brassica
7	174.8	6.1	2853	6	A97606	A97606 Sequence 3
8	174	6.1	4601	3	DMU11584	U11584 Drosophila
9	174	6.1	19517	3	DMU37541	U37541 Drosophila
10	172.4	6.0	67970	3	PFMAL1P3	AL031746 Plasmodiu
11	171.4	6.0	104992	2	AC005504	AC005504 Plasmodiu
12	171.4	6.0	169546	2	AC004157	AC004157 Plasmodiu
13	169.6	5.9	172816	2	AC093899	AC093899 Homo sapi
14	168.4	5.9	14867	3	AE001398	AE001398 Plasmodiu
15	167.8	5.9	172816	3	AC093899	AC093899 Homo sapi
16	167.2	5.9	86827	3	PFMAL1P5	AL034556 Plasmodiu
17	167	5.9	181792	9	AC098822	AC098822 Homo sapi
18	166.4	5.8	126999	9	AL513328	AL513328 Human DNA
19	165.2	5.8	104992	2	AC005504	AC005504 Plasmodiu
20	165.2	5.8	169546	2	AC004157	AC004157 Plasmodiu
21	164.8	5.8	14867	3	AE001398	AE001398 Plasmodiu
22	164.6	5.8	67970	3	PFMAL1P3	AL031746 Plasmodiu
23	164.4	5.8	159618	9	AC096750	AC096750 Homo sapi
24	161.2	5.7	86827	3	PFMAL1P5	AL034556 Plasmodiu
25	160.8	5.6	326924	2	AC093082	AC093082 Homo sapi
26	158.8	5.6	326924	2	AC093082	AC093082 Homo sapi
27	158.4	5.6	155106	2	AC104069	AC104069 Homo sapi
28	158.2	5.5	4601	3	DMU11584	U11584 Drosophila
29	158.2	5.5	19517	3	DMU37541	U37541 Drosophila
30	158	5.5	135250	9	AC104320	AC104320 Homo sapi
31	157.6	5.5	349980	6	AX344555	AX344555 Sequence
32	157.4	5.5	183584	9	AC012492	AC012492 Homo sapi
33	156.8	5.5	155106	2	AC104069	AC104069 Homo sapi
34	156.8	5.5	158398	2	AC011146	AC011146 Homo sapi
35	156.8	5.5	173915	9	AP001254	AP001254 Homo sapi
36	156.8	5.5	340000	9	AP001678	AP001678 Homo sapi
37	156	5.5	172724	2	AC098590	AC098590 Homo sapi
38	155.4	5.4	176174	2	AC007483	AC007483 Homo sapi
39	155.4	5.4	246080	2	AC104073	AC104073 Homo sapi
40	154.2	5.4	96953	2	AC079621	AC079621 Homo sapi
41	153.8	5.4	170143	2	AC048384	AC048384 Homo sapi
42	153.8	5.4	205130	2	AC105425	AC105425 Homo sapi
43	153.2	5.4	143331	9	AC091214	AC091214 Homo sapi
44	153	5.4	181792	9	AC098822	AC098822 Homo sapi
45	153	5.4	205130	2	AC105425	AC105425 Homo sapi

ALIGNMENTS

RESULT	1	
LOCUS	A97606	2853 bp
DEFINITION	Sequence 3 from Patent WO9915678.	DNA
ACCESSION	A97606	linear
VERSION	A97606.1 GI:6780907	PAT 26-JAN-2000
KEYWORDS		
SOURCE	unidentified.	
ORGANISM	unclassified.	
REFERENCE	1 (bases 1 to 2853)	
AUTHORS	Pelletier, G. and Drouaud, J.	
TITLE	MICROSPORE-SPECIFIC PROMOTER AND METHOD FOR OBTAINING HYBRID PLANTS	
JOURNAL	Patent: WO 9915678-A 3 01-APR-1999.	
FEATURES	AGRONOMIQUE INST NAT RECH (FR): PELLETIER GEORGES (FR)	
source	1. 2853	
BASE COUNT	961 a 506 c 381 g 996 t 9 others	
ORIGIN	/organism="unidentified"	
	/db_xref="taxon:32644"	

```
Query Match      . 99.7%; Score: 2844; DB 6; Length 2853;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2853; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

QY	1	gattccccaagaagaaacccgaagaagcaaatgttcccaacttcataataatataatttgt	60
Db	1	GGATTCACCAAGAGAAACCCAGAAAGCAAAATGTTTCTACCTTCATCAATATATATTTGT	60
QY	61	ttcagcctcatcatgttacaanaaccccttagctcaatggtatataatgtgttttag	120
Db	61	TTTCAGCTCATCATATGTATACAAACATCTCTTGTAGCTCAATGATATTAATGTGTGTAG	120
QY	121	atttcaataacccgggttcgaagcatagaacttgacacttttcacaacttttaaaagtgt	180
Db	121	ATTTCATTAACCCGGGTTGAGTCATAGACTTTGACACTTTTTCACACTTTTAAAAAGTGG	180
QY	181	aacgcacatacgcctgcagctgttcgacacggagatgtgnaactgtctcatataatgtag	240
Db	181	AACGCACATATCGCTGACGTGTGCGATCGAGAGTGATGCACTGCTATATATATGTAG	240
QY	241	attlaaagaaggagaccacagatgcgcgagtcgcacgagagatgcacactgcaat	300
Db	241	ATTTAAAGAGGAGAACCCAGCATATCGCTGACGCTGTGCAATCGAGATGTATGCACTGCCAT	300
QY	301	attataagctagatttgacgttaattccctttttaaactcttaataataataccagacttt	360
Db	301	ATTATATAGCTAGATTGTGACGTTATTCCTTTTAAATCTTAATTAATTAATACCAAGGCTTTT	360
QY	361	acttactaaatttmgcagctngtcatcagatgttatagmccttttttttganccgttga	420
Db	361	ACTTATTAATTTTGNGCATNGTTATCATCATGTTTATGCMCTCTTTTGTTCANCCGTTTGA	420
QY	421	tttgatttagcttatttgaatgtgncmagtlaagaatgaagaacaattatatttga	480
Db	421	TTTGATTATAGCTTATTGTAATGTAATGTCNCCNACCTAAGAAATGAGAACTAATTAATTGGA	480
QY	481	gaaatatataatlaatatgttcaataataatagaagaataatataatncccttggtttactgt	540
Db	481	GAAATATATTAATTAATATGTTCAATATATTAAGAAATATTATATNCCCTTGATGTACTGTA	540
QY	541	tggatggcgaagtgaagaactcttggaaataatatttgaagaactgaccttttccaaagaatga	600
Db	541	TGGATGGCGAGTGAAGATCTTTGTAATTAATTTTGAAGAACTTGCTTTTCCAAAAAGTGA	600
QY	601	aatatltgatalgttaacttaagttaacacatgaaatlaaataaataataataataata	660
Db	601	AATATTGATATGTATTAATTAAGTTAATACACATGAATAATTAAAAAAAATTAAATCAAAAAA	660
QY	661	gaaaaaactgtagtgaataccttcaagcttgaactatcttgygttccccccta	720
Db	661	GAAAAAAGCTGATAGATGATCTACCTTAACGCTTTTGAACTTAATCTGTGGTCAACCCCTTA	720
QY	721	aaccttaagttccccaacaataaattcaatatatgcatatctcatatccttttagaa	780
Db	721	AACCTTAAGTTCCACCAACAATTAATTTCAATTGCGATATCTATATCTTTTGAAA	780
QY	781	gtgaaacaaaatattatcaagttatatatttgttttccaataaagaataaanaaat	840
Db	781	GTGAACCAAAATATTATTCAGATTATATTATGTTTTTCAATAAAAAAGATAAAAAATTAAT	840
QY	841	aaaaataatagtagttacaaaaaanaaataataatatttaccagcgttcanaaacc	900
Db	841	AAAAAATATATAGTGTATCAAAAAAAAATTAATATTTTATACAGCGCTCAAAAAAC	900
QY	901	taaaccttaaaccttaatatattaacttttaggttaaaccttaaacctttgataaactt	960
Db	901	TAAACCTTAACCTTAATATATTAATCTTTTAGGTAAACCTTAACCTTTGGATTAATCTT	960
QY	961	aaacattaaacatlaaacaactlaaaccttaaccccttaaaccttaaaccttaagtgta	1020
Db	961	AAACATTAAACATTAAAAAACCTTAACCCCTTAATATCCTAACCTTAACCTTTAAGTGTTTA	1020
QY	1021	aatgtttaggttttttgattataagtttaggattataccaaagttttagttaccaca	1080

Db	1021	AATGTTAGTGTGTTTGGATTTATAGTTTAGATTTATTCGAAGGTTTAAGTTTACC	106
Qy	1081	gagcttctggttagaggatatacttagatttagtcttctaactgcagcgttcaagt	114
Db	1081	GAGTTTATGTTTAGGAGATATACATAGAGATTTACTGTTTACAGACAGCTCAAGT	114
Qy	1141	attttttaaaaaatatttttttgcataacactactatttttaattttttttcccttt	120
Db	1141	ATTTTAAAAAATATTTTTTTTGTAAACAACATATTTTTATTTATTTTAACTCTT	120
Qy	1201	tattttaaacaataataataatttaactccatcctgttcaataaagtgatgtaa	126
Db	1201	TATTTAAAAACATATATATATTTTAACTCCATCTGTTTCATTTAAGTGTCAATGAA	126
Qy	1261	catatttttttttatacaaaaaatgtgtaactttgaattccaatgcgaattattat	132
Db	1261	CATATTTTTTTTGTTCAAAAAATGTGTCACTTTGAAATTCGAATGCMAATTTATTA	132
Qy	1321	ttttcagctaaataatattgcgaagtgcatacttataaataatttattatcccaa	138
Db	1321	TTTTCCAGCTAAATATATGCAAAAGCGCATGTGCTATTAATATTTTATTTATCCCA	138
Qy	1381	atgcataatggcgaacacgtgtaataatagaacaactatattatattcaattttt	144
Db	1381	ATGCTATATTTGGCACAACACTGTGTTATTAATGAACATTAATATTTTATTTATTTT	144
Qy	1441	tcttaactctgtgtaaaaaatgtaaaagttaaatttattagaacgaattgataaatt	150
Db	1441	TCTTAATCTGTGTAATAAATGTCAAAGCTAAATTTATTAGAACGATGTAGTAATTT	150
Qy	1501	tgttcaattttttaagaatatacgataatagaataacacaatttattgtatgaagac	156
Db	1501	TGTTTATTTTTTAAAAAGATATCGAATATGGAATAAACAATTTATTTGTATGATGAAC	156
Qy	1561	taaaaaatcactocataagaagtggaacgcaagaataaagtaacgttttggggaagctaac	162
Db	1561	TAAAAATTCATCTCTAAGAGGTAAGCGAACAATAAATGTCACGTTTGGGGAAGCTAAC	162
Qy	1621	tatggcccaagaatgtaaaaaatcttctctgtattatacaaaatccttacaatttagta	168
Db	1621	TATGGCCCAAGTCAATCAAAATCTTCTCTGTATTTATCAAAATCTTACAAATTTAGTAA	168
Qy	1681	gagttaatagaccaacacatgatatacatatataagaaatcttcaaaaaattactagc	174
Db	1681	GAGTTAATAGCAACACATGATATATCATATTTAGAAATTTCTPAAAAATTTCTTAAC	174
Qy	1741	gaataataaaatcttcttatttatttatacaaaatccttataaaactttattatataac	180
Db	1741	GAAATATTAATAATCTTCTTTTATTTATCAAAATCTTTAANAACCTTTATATATATAC	180
Qy	1801	taaaacaattttaattaaagaanaataataggaccaatgagatatacaataatagatatt	186
Db	1801	TAAACAATTTTATATTAATAAATAATVAGGACCATATGATACATATAAATAATATATGTTAT	186
Qy	1861	tcttaagaatgataataatataataataacagttcatatattatcaaaataaataata	192
Db	1861	TCTTAAAGATAGTGAATATTAATATATATACCAATCATATATTTATCAAAAATAAATAATA	192
Qy	1921	tttttcgtagtcggaataatcaactataaaatcataaaacacacatgtagatgataatt	198
Db	1921	TTTTTCGTAGTCGCAATATCATATTAATAATTTATTAANAACAATATVAGATATATTT	198
Qy	1981	tattataataataataataaacccaagaagccttaccactgtatataacatacaaaacttt	204
Db	1981	TATTTATATATATATATATTAATAACCTTATAGCCCTTACCATCTGATATACATAAAACTTT	204
Qy	2041	ctctcgttttcgcttaactcaagcttcggaagaataaaaaacaaatgaagatgcaaac	210
Db	2041	CTTCTGTTTTCGCTTAACCTCAAGCTTCGAAGAATAAATAAACAATGAAGATGTCAAC	210
Qy	2101	ttgtctctgctatagatccctcttctaagctgtgtcaacatccaagaattacagcaacgaac	216

D	601	AATATTGATTTGTAACCTTAAGTTAAACAATGAAATTTAAAAAATTTAAATCAAAATA	660
Q	661	gaaanaacttgtagtgcctacccttcaacggtttgaactattccttggttcacccctc	720
D	661	GAANAACCTGATATGATGATTCACCTTCACACGTTTTCGAATATTCCTTGCTGCACCCCTA	720
Q	721	aaccttaagttcccaacaataaatttcattcttgcatctcattctcattctttagaa	780
D	721	AACCTTAAGTTCACCAACAATAAATTTCTATTCGATATCTCATCTTTTACAAA	780
Q	781	gtgaacaaaatttatacaagtatactgcttttccaataaagaataaataaat	840
D	781	GTGAACAACAAATATTTACAGGTTATATTTGTTTTTCAAATTAATAAAGATAAAAATTAAT	840
Q	841	aaaaataatagtagtaccaaaagaanaataatattttaccacgctcnaaaaac	900
D	841	AAAAAATAATAGTAGTTACAAAAAATAATTAATATTTTACCACGTCACAAAAAC	900
Q	901	taaacctaacccttaataattaaactttagtaacccttaacctttgataaactc	960
D	901	TAAACCTTAACCCCTAAATTAATTAACCTTTAGGTAAACCCTAACCTTGSGATTAATCTT	960
Q	961	aaacattnaacatctaaacaacaaacctaactcctaactcnaacccitgaagtta	1020
D	961	AAACATTAACCATTTAAACACTAAACCCCTAAATCCTTAACCTTAAGTGTTA	1020
Q	1021	aatggttagtgcttttgatlltagtcttaggattaccagaagtttaagtttaccoc	1080
D	1021	AAATGTTAGGTTTGTGATTTATGATTATAGGATTTATCCAAAGGTTTAAGGTTTACCCA	1080
Q	1081	gaattatggttttaggatttaccttaggttagtgctttacgcagacgttccaa	1140
D	1081	GATTTATGCTTTAGGATTAAGCTTAGGATTAAGTTTACTGACACACGCTTAAGT	1140
Q	1141	attttttaaasaatattttttttllaacaaactactatttttatttttttttaccctt	1200
D	1141	ATTTTTTTAAAAAATTTTTTTGTTCACAACTACTATTTTATTTTTTTACTCTTT	1200
Q	1201	tatattnaaacataataatttaaatcccatcgtttcataltgaaltgaaltgtaa	1260
D	1201	TATATTAAAAACATATATTAATTTAAATAGTCACATCGTTTCATATTAATTAAGTCAATGTAA	1260
Q	1261	catlacttttttggtacaaaanaatgctcactttagaatccaatgcaaaattttat	1320
D	1261	CATTTATTTTGTGTACAAAAAATGTCACCTTACAAATTCACATGCAAAATTAATTAAT	1320
Q	1321	ttttcagctaaatattatgcaaaagtcattgctcttataaataatttatllatctcaa	1380
D	1321	TTTTTACGTAATAAATTAATGCAAAAGTCATGATCTTAATAATTAATTTATTTATTCGCA	1380
Q	1381	atgcatacttggtccaacaatggttlaattatagaacctaataattatcttattctt	1440
D	1381	ATGCTATATTTGCTCAACACATGTGAATTAATGMAAACCTTAATTAATTTTCAATTAATTTT	1440
Q	1441	tccttaactcgtgtaaaaaagtcaaaglaaatttattagaacgaattgagtaattt	1500
D	1441	TCTTAATCTGTGTAAAAATGTCAAAGTAAATTTATTAGAAAGCAATGTAGTAATTTT	1500
Q	1501	tgtttcatttttttaagaagatcgaatttgaataacacaaatttattgtatgaac	1560
D	1501	TGTTTCATTTTAAAGATATGCAATATGAATATACCAATTTTATTTGATGATGCAAC	1560
Q	1561	taaaaatcactcobaagaagtgtaacgcagaagataagtaagtaacgctttggggaagctaac	1620
D	1561	TAAAAATTCATCTTAAGAAAGGTGAACGCAAGATATAGTCACAGCTTTGGGGAAGCTAAC	1620
Q	1621	tatggcccaagatcaacaacttttctgtatttatacaaaaaccttaacaaattagtia	1680
D	1621	TATGGCCCAAGATCAAAAATCTTTCTGTATTTATCAAAAATCCTTAACAAATTTTGTTA	1680
Q	1681	gagttaatagaccaaaacatgatattcatcattagaactttcnaaaatttactaac	1740

Db	1681	GAGTTAATAGACCAACACGATATACATCATTTAGATATCTGTAAAAAATTAAGTAC	174
QY	1741	gaataataaactcttcttattatccaanaactcttiaaaaactattatataac	180
Db	1741	GAAATAAATTAATCTTCTTTTATTTATCCAAATCCTTATTAANAACCTTATATATATAC	180
QY	1801	taaaacaatttaataaagaataaaggaccatggataataataatattgtattc	186
Db	1801	TAAACATTTTATTTAAAGAAATTAAGGTACATGAGTACATATAAAATTAAGTTATT	186
QY	1861	tcttaagtagtgaataataatataataccagtcataatttatacaaaataatata	192
Db	1861	TCCTAAGATAGTGTATATATTAATATATATACAGTCATATATTTATCAAAATAATATATA	192
QY	1921	tttttcgaagccgatatcatattactataaattctctaaacccatgtatgtatatt	198
Db	1921	TTTTTCGATGCCATATATATTACTATATAATTCATAAAACCATATGATATATATT	198
QY	1981	tattatataatataataaaccataacgcttaccactcgataacatcaaaactttt	204
Db	1981	TATTTAATATATATATATATTAATTAACCCCTTAACCTGTGATATACCATCAAAACCTTT	204
QY	2041	cttctcgcttcgctaactcaaggcttcgaaaagtaaaaaacatagaagatgcacac	210
Db	2041	CTTCTCGTTTCGCTAAGCTCAAGGCTTCCAAAAGTAAAAAAAACAATGAAATGTCACAC	210
QY	2101	ttgtctcttgatgacccctctttaaqtgtgtacatccaaagtttaagcaagcaaac	216
Db	2101	TTGTTCTTGCTATGATGCTCTTTTAACTGTGTACATCCAAAGTTACGCAACGAAC	216
QY	2161	taagatcatcaactaaccaagagctctctcatcgcggaacttacctgcgtttcaaccca	222
Db	2161	TAGATCATCACTAAACCAAGAGCTTTCATATGCGCGCACTTAACCTCGCTTCAACCCA	222
QY	2221	agcaaatitggccgttcgcgtgtctccggaaaagccttcacctgaagccacttcgacaa	228
Db	2221	AGCAAAATITGGCCGTTCCGCTGCTCGGAAAAGCCTTCCCTGCAAGCCACTTCGACTGA	228
QY	2281	ctcgcgtcatcttgcacaaaggaagtccaccagaigtgtgaacgaaaggaagtga	234
Db	2281	CTCGGTTCCATCTGCACAAAGAAAGTCACACATGCTTGAACGACAAGAGAGGATGAGTA	234
QY	2341	catgttttaataatatacgcgtagaacttcttcaaccaggaaagccgtattgtatcgaa	240
Db	2341	CATGTTTAAATGATATGCTGTAGACTTTCTTCAACGAGAAAGCCGTATATGATGGAAT	240
QY	2401	gtttgcgcgcgaatcaagaagatlgaaacaagaattgtlagaagaacgcttltgtatcttc	246
Db	2401	GTTTGCCTCCGCAATCAAGAAAGATGAAACAAAGATTGTGAGAAAGCCGTTTGGATCTTCC	246
QY	2461	atgaacccctcttgaacggcgatgttaagtlataattgtctccacgcgttgttgatctact	252
Db	2461	ATGACCCCTCTTGTGACCGGCTATGTCAAGGTACATATGTGCACCCCTTGTGATCTACTT	252
QY	2521	gaactctctctcaacaggtctcctttaaagtgtccttcttcaaggctcttcaaatgctc	258
Db	2521	GACCTCTCTTTCACAAGGCTCTTTACATGCTCTTCTTCAAGGCTCTTCAACATGCTC	258
QY	2581	cttcaacatgaccttcaacaggtccctttaaagtgtcctttaaigtcttcttataatgctc	264
Db	2581	CTTCAACATGCTCTTTCACAAGGCTCTTTAAAGTGTCTTTAAATGCTCTTTACATGCTC	264
QY	2641	cttataatgctccttcaacagcccttcaacagcccttcaacagcccttcaatgctc	270
Db	2641	CTTTACATGCTCTTTCACAAGGCCCTTTCACAGGCCCTTTCACATGCTC	270
QY	2701	cttactggccctcttgagaggtcccttcaacggcgtcaatgtattagctattatagaat	276
Db	2701	CTTACTGCCCCCTTGCAAGGCTCTTTCACCGGCTCACTGATTTAGCTATTTATGATGAAT	276
QY	2761	tattcaaglatatgatctcaaggaatttttttttctgttttaaaatttctgt	282
Db	2761	TATTCAGATATGATGCTCCAGGAGCTTTAACTTTTCTGTGTTTAAAAATTTCTGTT	282

QY 2821 tatttggagaaacgcttttgatttaactt 2853
 |||||
 Db 2821 TATTTGAGAAACCGCTTTTGATTTTAATT 2853

RESULT 3
 A97605 A97605 674 bp DNA linear PAT 26-JAN-2000
 LOCUS Sequence 2 from Patent WO915678.
 ACCESSION A97605
 VERSION A97605.1 GI:6780906
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.

REFERENCE 1 (bases 1 to 674)
 AUTHORS Pelletier,G. and Drouaud,J.
 TITLE MICROSPORE-SPECIFIC PROMOTER AND METHOD FOR OBTAINING HYBRID PLANTS
 JOURNAL Patent: WO 915678-A 2 01-APR-1999;
 AGRONOMIQUE INST NAT RECH (FR); PELLETIER GEORGES (FR)
 FEATURES Location/Qualifiers
 source 1..674
 /db_xref="taxon:32644"

BASE COUNT 155 a 181 c 124 g 214 t
 ORIGIN

Query Match 16.2%; Score 462; DB 6; Length 674;
 Best Local Similarity 81.9%; Pred. No. 36-53;
 Matches 614; Conservative 0; Mismatches 35; Indels 101; Gaps 3;

QY 2105 tcttgatgatctcttcttcttgaagcgtgtacatccaaagttaacgcaacagactaga 2164
 |||||
 Db 1 TCTTGATGATTTTCTTATGATGATGTCACATCCAAAGTCACAGCAACAGACTAGA 60

QY 2165 gtcatcaataacaaagagctctctcatcgcgcaacttaacctgttcaacccaaga 2224
 |||||
 Db 61 GTATCACTAACCAAGAGCTCTTCTATCGCGGACTTGCTCGCTTTCACCCCAAGCC 120

QY 2225 acattggcgttccgtgtgctccggaagacgtccctgcagcagcacttccgaactaacc 2284
 |||||
 Db 121 ACATTTGGCGGCTTGTGGGTCGGAAGAGCTTCCCTGGAGGCGACTTCCGACCAACTGC 180

QY 2285 gtccatctgcagcaaggaagtaacagatgttgaaagcaagaaggaaggaagtaacatg 2244
 |||||
 Db 181 GTTCCATCTGACACAGAGATGACAGATGCTTGTCCGACAGAGAGAGTATGATCATG 240

QY 2345 tttaatgatagcgtgagagcttcttccacgaagacgctattgagatcggaatgtg 2404
 |||||
 Db 241 TTTTGTATGATATGTTGAGACTTTCTTACACAGAGAAAGCGTTATTTGATGGAATGTTG 300

QY 2405 cgcgcgatacaagaagatgaacaagaatgtgaaagacgcttcttgatcttcaatga 2464
 |||||
 Db 301 CGCGGAGATCAACAGATGAAACAAAGATTGTGAGAGACGCTTTGATCTTTCATGA 360

QY 2465 cccctcttgacggtgtatgtcaagtaacatgtgtccacggttggatcttcaacc 2524
 |||||
 Db 361 CCCCTTCTTGACAGGCTATGTAACACTTCTCCACCGTGTGAGATCACTTCCAC 420

QY 2525 tctctctcaagcgtcttcaatgatgtcttcttcaagcgtcttcaatgatgtctctc 2584
 |||||
 Db 421 TCTCTCTTC----- 429

QY 2585 acatgtccttcaagcgtcttcaatgatgtcttcaatgatgtccttcaatgatgtcctt 2644
 |||||
 Db 430 -----ACATGCTCTCTTC 441

QY 2645 acatgtccttcaagcgtcttcaatgatgtccttcaatgatgtccttcaatgatgtcctt 2704
 |||||
 Db 442 ACAGGCTCTTACATGCTCTCTTACAGAGGCTCTCTTACATGAGCCCTTACAGAGGCTCTCT 501

QY 2705 actgcccccttcgagcgtcttccacggtctcaatga-ttaagctattgatagaattat 2763
 |||||
 Db 502 ACTGCCCGCTTACAGAGCTCTCCACGGCTGATGATTTTACCTATTTGTGAATTTAT 561

QY 2764 tcaagatgtatgtctcgaagagtttagtttttcttctgttttaaaatttggttat 2823
 |||||
 Db 562 TCAAGTGTGATGCTCTTACAGAGGAGTTTATG-GTTTTCTTGTTTAAATTTGTGTTAT 620

QY 2824 tttagaagaaacgcttttgatttaactt 2853
 |||||
 Db 621 TTGAGAAACCGCTTTTGATCTTAACTT 650

RESULT 4
 A97604 A97604 497 bp DNA linear PAT 26-JAN-2000
 LOCUS Sequence 1 from Patent WO915678.
 ACCESSION A97604
 VERSION A97604.1 GI:6780905
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.

REFERENCE 1 (bases 1 to 497)
 AUTHORS Pelletier,G. and Drouaud,J.
 TITLE MICROSPORE-SPECIFIC PROMOTER AND METHOD FOR OBTAINING HYBRID PLANTS
 JOURNAL Patent: WO 915678-A 1 01-APR-1999;
 AGRONOMIQUE INST NAT RECH (FR); PELLETIER GEORGES (FR)
 FEATURES Location/Qualifiers
 source 1..497
 /db_xref="taxon:32644"

BASE COUNT 107 a 130 c 70 g 190 t
 ORIGIN

Query Match 13.8%; Score 393.2; DB 6; Length 497;
 Best Local Similarity 98.0%; Pred. No. 5.8e-44;
 Matches 398; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2448 tttagatcttccatgagcccttcttgaccggtatgtcaagctatgtccacgctt 2507
 |||||
 Db 1 TTGGATCTTTCCATGACCCCTTCTTGACCGGCTATGTCAAGTACATGCTCCACCGTT 60

QY 2508 gtggatctacttcaactctctcttcaagcgtcttcaatgatgtccttcttcaagcgt 2567
 |||||
 Db 61 GTTGGATCTTACCTTACCTCTCTTCAAGGCTCTTACATGCTCTTCTTCAACAGGCT 120

QY 2568 ccttcaatgctcttccatagctcttccacagcgtccttcaatgatgtccttcaatgct 2627
 |||||
 Db 121 CTTTACATGCTCTCTTACATGCTCTTCAACAGGCTCTTCAATGCTCTTTAAATGCT 180

QY 2628 ccttcaatgctcttccatagctcttccacagcgtccttcaatgatgtccttcaatgct 2687
 |||||
 Db 181 CTTTACATGCTCTCTTACATGCTCTTCAACAGGCTCTTCAACAGGCTCTTCAACAGGCT 240

QY 2688 ccttcaatgctcttccatagctcttccacagcgtccttcaatgatgtccttcaatgct 2747
 |||||
 Db 241 CTTTACATGCTCTCTTACATGCTCTTCAACAGGCTCTTCAACAGGCTCTTCAACAGGCT 300

QY 2748 tattgataaataatcaagatagctcgaagagtttagtttttcttctgtttt 2807
 |||||
 Db 301 TATTTGATGATATTTACAGAGTATGATGCTTGAAGAGTTTATGATTTTCTGTGTTT 360

QY 2808 aaaaatttgtttatttttgagaagacgcttcttgatcttcaact 2853
 |||||
 Db 361 AAAGTTTGTGTTTATTTTGAAGAAACCGCTTTGATTTTAATT 406

RESULT 5
 AB005238 AB005238 75188 bp DNA linear PLN 27-DEC-2000
 LOCUS
 DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, p1 clone:MKP11.

ACCESSION	AB005238	BA000015
VERSION	AB005238.1	GI:22643310
KEYWORDS		
SOURCE	Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1 clone:MKP11.	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (sites)	
AUTHORS	Sato,S., Kotani,H., Nakamura,Y., Kaneko,T., Asamizu,E., Fukami,M., Miyajima,N. and Tabata,S.	
TITLE	Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence features of the 1.6 Mb regions covered by twenty physically assigned P1 clones	
JOURNAL	DNA Res. 4 (3), 215-230 (1997)	
MEDLINE	97471969	
REFERENCE	2 (bases 1 to 75188)	
AUTHORS	Nakamura,Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (02-JUL-1997) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynakam@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934)	
COMMENT	Address for correspondence: kaos@kazusa.or.jp For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MKP11 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Graal (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/Graal-1.3/), GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brenkel, Stanford University, http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is F2K13 and the 3' clone is T10B6.	
FEATURES	Location/Qualifiers	
source	1. 75188	
	/organism="Arabidopsis thaliana"	
	/strain="Columbia"	
	/db_xref="taxon:3702"	
	/chromosome="5"	
	/clone="MKP11"	
	/clone_lib="Mitsui P1"	
	complement(1..99)	
exon	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=3	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(185..680)	
exon	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=2	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
exon	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	
	gb AA26111.1	
	gene_id:F2K13.310"	
	/number=1	
	/evidence="not experimental	
	/product="protein; similar to unknown protein"	
	complement(1306..1419)	
	/note="CDS is reported in Acc# AP002072	

Qy	1019	taaatgcttggttttgcattatagtttaggaattacccaagglttaagttacc	1078
Dd	990	TAGGCTTACTGTATTAATGTTAAAGTTAATAATTATCCAAAGGTTTAGGTTTACCT	931
Qy	1079	aagaattacgcttagggagcttagacttagtgattgcttagtcagcagctcaaa	1138
Dd	930	AAAAGTTTAAATATTTTGCGTTTAGGTTTAA-----TCGTTTTGACGGCTGTAAAA	879
Qy	1139	gtattttttaaaaaatatlttttgttaacaactactattttattttttacct	1198
Dd	878	ATATT-----AAATTTTTTTTTTTTTTAACTACTATATTTTTTATTTTATCTT	825
Qy	1199	tttataattaaacaatacataaatttaacct	1230
Dd	824	TTTATTTGA AAAACAATATATTAAGTTGATAT 793	
RESULT	8		
LOCUS	DM011584/C	4601 bp	DNA linear INV 23-JUL-1994
DEFINITION	DM011584	Drosophila melanogaster Oregon-R mitochondrial A+T region:	
ACCESSION	U11584	U11584.1 GI:508826	
VERSION	U11584.1	GI:508826	
KEYWORDS	mitochondrial DNA; A+T region; tandem repeats.		
SOURCE	fruit fly		
ORGANISM	Mitochondrion Drosophila melanogaster		
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
AUTHORS	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;		
TITLE	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
JOURN	Lewis,D.L., Farr,C.L., Farquhar,A.L. and Kaguni,L.S.		
MEDLINE	Sequence, Organization and Evolution of the A+T Region of		
REFERENCE	Drosophila melanogaster Mitochondrial DNA		
AUTHORS	Mol. Biol. Evol. 11, 523-538 (1994)		
JOURNAL	2 (bases 1 to 4601)		
TITLE	Kaguni,L.S.		
AUTHORS	Direct Submission		
JOURNAL	Submitted (28-JUN-1994) Laurie S. Kaguni Ph.D. Dept. of		
FEATURES	Biochemistry, Michigan State University, East Lansing, MI,		
source	44824-1318, USA		
Location/Qualifiers			
I. 4601			
/organism="Drosophila melanogaster"			
/strain="Oregon-R"			
/db_xref="taxon:7227"			
/dev_stage="embryo"			
650..1022			
/note="repeat I-A"			
/rpt_type=tandem			
1023..1360			
/note="repeat I-B1"			
/rpt_type=tandem			
1361..1705			
/note="repeat I-C/A"			
/rpt_type=tandem			
1706..2043			
/note="repeat I-B2"			
/rpt_type=tandem			
2044..2388			
/note="repeat I-C"			
/rpt_type=tandem			
2491..2511			
/note="deoxythymidylate stretch"			
2512..2648			
/partial			
/rpt_type=tandem			
2649..3112			
/note="repeat II-A"			
/rpt_type=tandem			
3113..3576			

```

      /note="repeat II-B1"
      /rpt_type= tandem
      repeat_unit 3577..4040
      /note="repeat II-B2"
      /rpt_type= tandem
      repeat_unit 4041..4504
      /note="repeat II-C"
      /rpt_type= tandem
      complement(4565..4585)
      misc_feature /note="deoxythymidylate stretch"
      BASE COUNT 2271 a 131 c 74 g 2125 t
      ORIGIN
Query Match 6.1%; Score 174; DB 3; Length 4601;
Best Local Similarity 47.7%; Pred. No. 1.1e-14;
Matches 815; Conservative 0; Mismatches 868; Indels 25; Gaps 10;
QY 299 atataacgtagattcagcttattccctttaaacttaataataacagngctt 358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3543 ATATTAATTTATTTAATATTTAATTTTCATATATATATATATATATATAT 3484
QY 359 ttaacttaattctgngcagctatcagctgcttgcctcttcttgcagcgtt 418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3483 CAATTAATATATTAATTTATTAATATTAATTAATTAATTAATTAATTAAT 3424
QY 419 gatggttatagcttatttgaaatgngcagctaaagaaagaaacattatcttg 478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3423 AATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3365
QY 479 gagaaataataatlaataatgctcaatataagagaataataatccttgatgctg 538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3364 TTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 3305
QY 539 taagatgagcagagaagatcttgcataatcttgagaactgcctcttccaaaagt 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3304 TATATAT---ATTAATATCTAATTAATTAATTAATTAATTAATTAATTAAT 3249
QY 599 aaaaatttgatagtaacttaagcttaacacatgaaataataaaaaaatlaatac 658
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3248 AATTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 3189
QY 659 tagaanaaactgtagtgcctaccccttcaacgcttgcgaacttcttgcacacccc 718
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3188 TATTAATTTATTTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTT 3129
QY 719 taaaccttaagctacacaaacataaaatcttcaatctgcatctctatctttaga 778
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3128 TAAATATTTT---TTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3072
QY 779 aagtgaaacaaatataatcaagcttataatgcttcttcaaatcaaaagaataaaata 838
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3071 TTTATTTAAATTTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 3012
QY 839 ataaaaataatagctgctacaaaaaaaaataatlaatttcaacagcgtcanaaaac 898
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3011 TATTAATTTTAAATTTAATTAATTTAATTAATTTAATTTAATTTAATTTAAT 2952
QY 899 actaaacaaacccataataatcaactttagtgtaaaccccaaccttgatataac 938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2951 TAAATTAATTAATTAATTAATTAATTAATTTT---TATTAATTAATTAAT 2893
QY 959 ttaaacatlaaacatlaaaacactaaacccataacccctaaacccctaaagctt 1018
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2892 TTTAAATTTTAAATTTTAAATTTAATTAATTAATTAATTAATTAATTAATTA 2833
QY 1019 taaatgcttagctgctttagctttagcttagcttagcttagcttagcttagcttag 1078
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2832 TAAATATCTTTTAAATTAATTAATTAATTAATTAATTTTAAATTAATTAATTT 2773
QY 1079 aagaagcttagcttagcttagcttagcttagcttagcttagcttagcttagcttag 1138
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2772 TTTAAATTTTAAATTTTAAATTTAATTAATTAATTAATTTTAAATTTAATTA 2718

```

```

QY 1139 gtaatttttaaaaaataatttttttgaacactactattttattatttttttacc 1198
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2717 TTTATTTATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 2658
QY 1199 ttatactaaacataataatlaatttaacccatctgcttcaatlaaagctgcatgt 1258
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2657 TTTTATTAATTAATTAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 2598
QY 1259 aacattattttttgt---tacaacaaatctgcatcttagaattccaatgcaaaat 1315
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2597 AATTTAATTTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2538
QY 1316 ttatttttcaagctaaattatgcaagctgcatctgacttaataataattattat 1375
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2537 AATATTAATTAATTAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 2478
QY 1376 ctcaaatgctatattgctcaacatgctgtaattatgaacatlaataattcat 1435
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2477 TTTAATTTAAATTAATTAATTAATTTAATTTAATTTAATTTAATTTAATTT 2418
QY 1436 tttttcttaactgctgtaaaaaatgctaaagtaaatatttagaagaagatgagtaa 1495
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2417 TATTAATTAATTAAT---TAAATTAATGAATTAATTAATTTAATTAATTAAT 2361
QY 1496 tattttgctcatctttttaaagaatcgaatcgaatgaataacaaacttattgcat 1555
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2360 TTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 2301
QY 1556 gaacctaaaaatcattccatcaagaagctgaaagcgaagaataagcttctggggaag 1615
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2300 ATTTTATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 2244
QY 1616 ctactatgcccccaagctcaaaaactcttctgtattatcaaaatcccttcaaat 1675
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2243 TTTATTAATTAATTAATTTAATTAATTAATTTTAAATTTTAAATTTTAAATTT 2184
QY 1676 agttagagcttaataagcaacacatgattatcatcattatgaatatctcaaaaa 1735
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2183 AATTTATTAATTAATTTAATTAATTAATTAATTAATTTAATTTAATTTAATTA 2124
QY 1736 ctacgaaataataaaactcttcttattatcaaaaacttataaaacttattata 1795
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2123 TATATATTAATTAATTAATTAATTTAATTTAATTTAATTTAATTTAATTTAAT 2064
QY 1796 tatactaaacaaatlaataataaa-gaaaataaggaacatctgatacataaaatata 1854
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2063 TATATCTTAATTAATTTAATTAATTAATTTAATTTAATTTAATTTAATTTAAT 2004
QY 1855 gtaattcttaagaagtagtaataatlaataataataacagccatataatttcaaaa 1914
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2003 TGAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTAAT 1944
QY 1915 ataatatttctgtagcagataatca-ttaataaaatcattcaaaaacacatgtagat 1973
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1943 TTTAAATTAATTTAATTTAATTAATTTAATTTAATTTAATTTAATTTAATTTAAT 1884
QY 1974 tataatttatttataatataatataaa 2001
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1883 AATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTAAT 1856

```

RESULT 9
 DM037541/c
 LOCUS 19517 bp DNA circular INV 02-MAR-2001
 DEFINITION Drosophila melanogaster complete mitochondrial genome.
 ACCESSION U37541
 VERSION 037541.1 GI:1166529
 KEYWORDS
 ORGANISM
 fruit fly
 Mitochondrion Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

REFERENCE Muscomorpha, Ephydroidea, Drosophilidae, Drosophila.
AUTHORS 1 (bases 12511 to 12682)
Clary, D.O., Goddard, J.M., Martin, S.C., Fauron, C.M. and
Wolstenholme, D.R.
TITLE Drosophila mitochondrial DNA: a novel gene order
JOURNAL Nucleic Acids Res. 10 (21), 6619-6637 (1982)
MEDLINE 83090428
REFERENCE 2 (bases 5269 to 5695).
AUTHORS Clary, D.O., Wahlthner, J.A. and Wolstenholme, D.R.
TITLE Transfer RNA genes in Drosophila mitochondrial DNA: related 5' flanking sequences and comparisons to mammalian mitochondrial tRNA genes
JOURNAL Nucleic Acids Res. 11 (8), 2411-2425 (1983)
MEDLINE 83220794
REFERENCE 3 (bases 404 to 5272)
AUTHORS de Bruijn, M.H.
TITLE Drosophila melanogaster mitochondrial DNA, a novel organization and genetic code
JOURNAL Nature 304 (5923), 234-241 (1983)
MEDLINE 83245048
REFERENCE 4 (bases 804 to 1778)
AUTHORS Satta, Y., Ishiwa, H. and Chigusa, S.I.
TITLE Analysis of nucleotide substitutions of mitochondrial DNAs in Drosophila melanogaster and its sibling species
JOURNAL Mol. Biol. Evol. 4 (6), 638-650 (1987)
MEDLINE 88174373
REFERENCE 5 (bases 5268 to 13619)
AUTHORS Garesse, R.
TITLE Drosophila melanogaster mitochondrial DNA: gene organization and evolutionary considerations
JOURNAL Genetics 118 (4), 649-663 (1988)
MEDLINE 88212147
REFERENCE 6 (bases 441 to 2967)
AUTHORS Satta, Y. and Takahata, N.
TITLE Evolution of Drosophila mitochondrial DNA and the history of the melanogaster subgroup
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9556-9562 (1990)
MEDLINE 91088557
REFERENCE 7 (bases 14215 to 14512)
AUTHORS Ballard, J.W., Olsen, G.J., Faltz, D.P., Odgers, W.A., Rowell, D.M. and Atkinson, P.W.
TITLE Evidence from 12S ribosomal RNA sequences that onychophorans are modified arthropods
JOURNAL Science 258 (5086), 1345-1348 (1992)
MEDLINE 93088057
REFERENCE 8 (bases 14917 to 19517)
AUTHORS Lewis, D.L., Farr, C.L., Farquhar, A.L., and Kaguni, L.S.
TITLE Sequence, organization, and evolution of the A+T region of Drosophila melanogaster mitochondrial DNA
JOURNAL Mol. Biol. Evol. 11 (3), 523-538 (1994)
MEDLINE 94285822
REFERENCE 9 (bases 1 to 408, 13319 to 19517)
AUTHORS Lewis, D.L., Farr, C.L., and Kaguni, L.S.
TITLE Drosophila melanogaster mitochondrial DNA: completion of the nucleotide sequence and evolutionary comparisons
JOURNAL Insect Mol. Biol. 4 (4), 263-278 (1995)
MEDLINE 96423163
REFERENCE 10 (bases 1 to 19517)
AUTHORS Lewis, D.L., Farr, C.L. and Kaguni, L.S.
TITLE Direct Submission
JOURNAL Submitted (03-OCT-1995) Laurie S. Kaguni, Biochemistry Department, Michigan State University, East Lansing, MI 48824-1319, USA
FEATURES
source location/Qualifiers
1. 19517
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/note="derived from new and previously submitted sequences; sequence is a composite containing sequences obtained from different Drosophila melanogaster strains"
1. 65
/product="tRNA-tle"
complement(97. 165)
tRNA
tRNA

tRNA
CDS
/product="tRNA-Gln"
171. 239
/product="tRNA-Phe"
240. 1265
/codon_start=1
/transl_table=5
/product="NADH dehydrogenase subunit 2"
/protein_id="AAC47811.1"
/db_xref="GI:1166530"
1264. 1329
/product="tRNA-Trp"
complement(1322. 1383)
/product="tRNA-Cys"
complement(1403. 1468)
/product="tRNA-Tyr"
join(1470. 1472,1474. 3009)
/note="mechanism underlying reading frame shift after first codon uncertain"
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit I"
/protein_id="AAC47812.2"
/db_xref="GI:7412849"
/translation="MSKWLSTNKHKIDGTYIFGAGAVGTSLSILIRAEIGHPG
ALIGDQIYNVITVAHAFIMFEWNPIMIGFGNMLVPLMGAPDAFPPNNNSFW
LILPALSLILSSWVENGAGTGVYPPPLAGIAGASVLAIFSLIAGISSILGA
VNEITTVINMRSTGISLDRMPLFWSVITALLLSLPLVAGATMTLDRNITSF
FPAGGDPILLYOLHFMFPGHEVYLLIPGCMSHIISDSKKEKPGSGIYYAM
LAIGLIGFTVMAHMFMTGMDVDFAYFTSNMIIAYTGKITSWALTIGTOLSTS
PALMAAGVPLFTVYGLGAGVLAASSVDIILHDTTYVAHFHYLSGAVAFVIMAG
IHMYLFTGLTINNNKMLKSHFLIMFTGVNLTFEPDHFGLAGMPRRSDVDAYTVN
IVTIGTSTLSLGLIFFEFTIWESLVSOROVITYPLNLSSTIEMVQNPABHSSELP
LITN"
3012. 3077
/product="tRNA-Leu"
3083. 3767
/note="TAA stop codon is completed by the addition of 3' A residues to the mRNA"
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit II"
/protein_id="AAC47813.1"
/db_xref="GI:1166532"
/translation="MSTVANGLDSDASPLMEOLIFPHDALLILMTIVLYGMEM
LFNNYVARELLHQLIEMITLIPAILLFIAPLSRLYLDELINPEVTLKSIGH
QWVSYEYSDNNNEIPDSYMPITMELMTDFRLDVRNRYVLPNNSQIRLVAAADY
HSWVVPALGVAVDGTPLRLNQTNFINRPGLEFGCCSEICGAMHSPMFIYESPVNY
FIKWSNNS"
3768. 3838
/product="tRNA-Lys"
3840. 3906
/product="tRNA-Asp"
3907. 4068
/codon_start=1
/transl_table=5
/product="ATPase 8"
/protein_id="AAC47814.1"
/db_xref="GI:1166533"
/translation="MPQMAPIWLLFLIFISITFLFCSINYSPNSPKSNEIKNI
NLNSNMKW"
4062. 4736
/codon_start=1
/transl_table=5
/product="ATPase 6"
/protein_id="AAC47815.1"

/db_xref="GI:1166534"
/translation="MATNLFSVDEPLAIFNFSIANLSTFLGLMIPSIYWLMPSRYNIMNSILITLHKKEFTLPGSGNGSTFIEISFLIENNGLEPPYIPTSTSLTLTSLALPMLCEMLYGMINHOHMFAMIVPOCPALILMPCVCIENISNIRGCTLAIVLTANAIAGHLITLLIGNTSSMSYMLMFLMAQIALVLVESAVAMIQSYVAIVLSTLXSEVN"
CDS
4736..5524
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit III"
/protein_id="AAC47816.1"
/db_xref="GI:1166535"
/translation="MSTHSNHPHLVDYSPMPDLCAGAMTTVSGNWKVPHOYDLSLVGNIITLTVYQMRDVSREGTYOGLHYAVTIGLRMGMLIFLSEVFEFPMAFHSSPAIEIGASNPMPGIIISFNPEOPLNLTALLASGVTVNAHSHLMNNSOTLOGTFVNLGIGFTFIIQAYEIEAPRIADSIYSTPEPMAGPCHILICTFLLVCLRLHNLNFKSNHHPFEBAAMVNHFPDVVFLFLITITVMWG"
trna
5543..5607
/product="tRNA-Gly"
5608..5961
/codon_start=1
/transl_table=5
/product="NADH dehydrogenase subunit 3"
/protein_id="AAC47817.1"
/db_xref="GI:1166536"
/translation="MESIIFALLILLITTVIMPLASILSKKALIDREKSSPECGDPKSSRLPESEFLITITIFLIFDEIALLIPMIIMKSNIMWTISIFILLILIGLYHMOGMLWMSN"
trna
5981..6045
/product="tRNA-Ala"
6055..6118
/product="tRNA-Arg"
6119..6183
/product="tRNA-Asn"
6184..6251
/product="tRNA-Ser"
6252..6318
/product="tRNA-Glu"
complement(6337..6401)
/product="tRNA-Phe"
complement(6401..8124)
/note="TAA stop codon is completed by the addition of 3' A residue to the mRNA"
/codon_start=1
/transl_except="(pos:complement(6401..6402),aa:TERM)"
Query Match 6.18; Score 174; DB 3; Length 19517;
Best Local Similarity 47.7%; Pred. No. 8.7e-15;
Matches 815; Conservative 0; Mismatches 868; Indels 25; Gaps 10;
QY 299 atataacgtagattgacgtatcttcctttaaacttaataataacagagctt 358
DB 18459 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 18400
QY 359 ttacttaataatttngcangtcatcagtggttaagcctcttcttttgcancgtt 418
DB 18399 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 18340
QY 419 gattggttaactcttgaatgctgcacgtaagaagaagaagaacttaacttg 478
DB 18339 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 18281
QY 479 gagaataataatttaataatgctcaatataagaagaataataatactgactg 538
DB 18280 TTTAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 18221
QY 539 taatgatacgagtagaagacttgaataatattggaacttgccttccaaaagt 538
DB 18220 TAT 18165
QY 599 aaaaatttgatactgaacttaagtaacacagaaataaaaaaataataataa 638
DB 18164 AAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 18105

QY 659 tagaaaaaacgtatgatacttaacccttaacgctttgaaactattcttgctaccccc 718
DB 18104 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 18045
QY 719 taaccttaagtcacccaacataaaatttcattatgcatattctatcctttaga 778
DB 18044 TAAATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 17988
QY 779 aagtgaaacaaaatattatcaagttatattatgcttttcaataaagaataaata 838
DB 17987 TTAATTAATTAATTAATTAATTTTCTATATATATATATATATATATATAT 17928
QY 839 ataaaaataatgctgtaacaaaaaaataaatttaccagcgctcanaaac 898
DB 17927 TATTAATTTTATATATATATATATATATATATATATATATATATATAT 17868
QY 899 actaaaaaccctaaacccataaacttttagtaaacccctaaaccttgataaac 958
DB 17867 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 17809
QY 959 ttaaacatttaacatttaaacacttaaaccttaaccttaaaccttaacgctt 1018
DB 17808 TTAATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 17749
QY 1019 taaatgcttagcttttgaattatagtttaggaattatccaaaggttaaggtta 1078
DB 17748 TAATATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 17689
QY 1079 aagaagttatgcttagggatcatgacttaggaattagtttactgacgacgtcaa 1138
DB 17688 TTAATTAATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 17634
QY 1139 gttatttttaaaaaataatttttttgaacaaactacatttattatttttacct 1198
DB 17633 TTAATTAATTAATTAATTTTGTATTTTAAATTAATTAATTAATTAATTAAT 17574
QY 1199 ttataataaacaataataataatttaactcatcctgcttcatatataatgctt 1258
DB 17573 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 17514
QY 1259 aacattatttttctgt--tacaaaaaaaltgcactttaggaattccaatgcaaa 1315
DB 17513 AATTAATTTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 17454
QY 1316 ttatttttcaagctaaataatattgcaagtgcatgacttataataatttattat 1375
DB 17453 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 17394
QY 1376 ctcaaatgctatattgctcaaacatgtaataatgaacttaattatattcat 1435
DB 17393 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 17334
QY 1436 ttcttcttaactcgtgttaaaaatgtaaaatgtaaaatlltagaaacgaatgtaa 1495
DB 17333 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 17277
QY 1496 tatttgccttcaatttttaaaagatatcgatatgaataacacaaatttactg 1555
DB 17276 TTTATTAATTAATTAATTAATTAATTTTCAATTTAAATTAATTAATTAATTA 17217
QY 1556 gaacctaaaatcatcctcaagaagtgcaagcgcaagaataagtaacgctttg 1615
DB 17216 ATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 17160
QY 1616 ctactatgcccagaagtcacaaatcttctgtatattatcaaaatctccaactt 1675
DB 17159 TTTATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 17100
QY 1676 agttaaggttaataagaccacaacacatgatatcatcatttagaataattc 1735
DB 17099 AATTAATTAATTAATTTATGAAATAGGGGAATTAATTAATTTTCAATTAATTA 17040

```

Protein: 18= "conserved hypothetical protein, UPF0006 family"/
/db_xref= "GI:5594244"/
/translation= "MKLVFHYIVLVFYIISFLKNSLKIYNDLVLEITVANKVKKV
LQIKRSMKKNKHNRIMEDNSKFIIDISNLIDKKFQDGVSKKENDLQNVLPNARK
NNMDDIITITCAIEDISLKECNGEYDPGKRLIYASVHPHTNVEYEDLNKHEEKE
IIAAKEYEELFKYKKNQDQNSKMGKNGKIKGDEMDNNILLENDDITPGFKIT
NKDEYEELENNKRIIVYPRVYICIGIDLPDRLAFCKRIQIQTIFLQIDVQNNP
LPMFLAHMNSSEPPKVIYDITKFLKENGSGVASHFDASVEYIIIVQNNYNNKLVG
CSKSLSEINAAVKKIPNLILITLDAIPKQGVKTHASVEYIDITYEAKVYNNKLVIN

```

```

/note="MAL1P3_03, putative ABC transporter, len: 1822 aa"
/codon_start=1
/product="putative ABC transporter"
/protein_id="CAB63558.1"
/db_xref="GI:6594246"
/training="NTYIKENVGJLNGKNGKKKSCONISFLNLSFDMIRPLINDLIK
GDIOLPLNCRNEDVPYPAKSLBEENLDVEDESEFYSKSNSENHVLAHNCNSDAE
KAYANVYINILMIKTEFRKILHLSFLEETLVLVLGGCFIDYARILEGQKIPV
YLSFLEGEVSGVLVVAIMEFFLEHFLFELTINLKLSTVLYTGLNCSNG
NHLONDAVDNYNRFESSQSEIDELSDPFLISGKNASSSGIKRNNKNNTEYEN
DYLINISKTKMEKDSLENBRSLPNVINYINISMDVSPYPTFSCINLNFVYKIFL
MSFYVHLKIGSNGVGLAINSLAYSALTEFFLPSLPSKSYLYLRYRKRIDNHHVL
KEKRLKKNPNESEAFKINIFPKREKCKCKIRLYLSNIGFISISDSIVDEYIFL
YLDNRKNKEELKFTSLIMPLYLILISVANGPILVANNWEGVINYIRLNYIND
LYVNDIKNFWYRTRYNEDYIVDVTGLFONENITSHDQSNHLYKANYTKNLTN
MKRYFFYKAKMYKHKNILNKOLISGLKLVNDNTKNTICPOEHSNRYNYNSHIE
KKEEYENHNSNSGSMSEPKRKKNNKYITLKNCSGLSGLSDKDCDNHDKLNTIN
LRNSLAITLGNSSGCKSAFPHSLDPMWNTGNYITIEFFRKAPLITLYQNSMLFG
NIRSMILGNEYNLITYITLLOSELNLDNLSTIEGDKRYINDNHNLSKQOYVILG
RALYVLAIHMRKLEDEKRLIQPNEILDKDILNNKNISYNNKSGKLVNINPENN
YLOKLMCDNNMYLTDLDFTSLDPSISKFLSFKCKENISFKONGSVITLISMS
LTDNPLLEDIDLVQYEVNIEFDQTKTLVGRNISVEWKNLNTITSEHWGVSNTL
IDYTIKRLIDENHVKHNSKMYKEAVFYGNRTSEYSEIDSLINKEITKRAKKNY
KKEHNNKNNNNNNNNNSKDDHININDNHRVNDIINDGPSVDETSYKILKAKKNY
YTLDTYISNSDKEIYKPLKQKHEEPKSSMPKVSNNKYNNSKRYEDNSNS
RGSISLETTYLMTFOYQVFLYVTFMILSIFDEKIEFVFLPMMSIISKNNHEISD
TLOQVYIELEFVLILISLVTSYGCIFMIFVNLGNTAIVHNHNLITLNLPLVIF
YNNNLGNINNEFLIDISAPYGLFKIRYAPFIPRCLISLTLIYIRRCITPIFV
ILYIEFEKFEKSSGEQRLYLSCHPLVCNITSNALSNGININITYKNTYHLYVE
YHINNRSITYFEKRLINRNASLYKLFLETLTLYTIMHPLVYASGLIKLYKENVYRI
LSTIGCISLARGLVITKFLCDVYTHIEKENCVCORLEPFAKISNKNENAKNEEL
NVITTYQKKEENISDKISALVEYKKNLSISINSQDDKESKYSIKRENVYYSK
KPTLVNGYIKYIDEEPSTKINIMYALKNOKIGLYSGDAKSKTILSLILGINISOG
KITYEGVDIRYNNKGEDSLITGLIASSQFVFNWNIREFIDPVANNPDEIDYALKLN
GILNGKNDILRYNKKODMSKYNKILIOSKYSVONSNPTLITNDCITRYLSYRLTN
BHKKILILDEIPNINNSYHDELNSFLIGKASFPNTIKNHPNNTYVILISHANT
LSCSDIYIYLRKGEITYRCSYEDVAKTOSLSHLEMD"
23896. 31533

```

gene subunit RNA genes and spacer regions
23896..31533
/gene="rRNA"
complement(join(31966..32476,32675..32775))
/gene="MALP3.04"
complement(join(31966..32476,32675..32775))
/gene="MALP3.04"
/note="MALP3.04, conserved hypothetical membrane protein, len: 203 aa, similarity: P. falciparum chromosome 2, PEB010W, O96126 predicted integral membrane protein (255 aa) fasta scores: opt: 335, E(): 4.9e-15, (36.18 identity in 191 aa overlap)"
/codon_start=1
/product="conserved hypothetical membrane protein, MALP3.04"
/protein_id="CAB63559.1"
/db_xref="GI:6594247"
/translation="MKSYTFLNVTILLFLFLFYVYNDFSTKFKNNIKIDIN RFRRIASEBOKYPWEDPCLINBELIRPENNDSPYLPHEHENDKINLSINS TKIMETIKKRONYKEDTNMNMMDPMHYKANTILYVHKILNITLTKDLNPI HDKERTITTKIWOEDIEYFLFNLOVEMLRILTLLEFYKKE"
complement(32477..32486)
/gene="MALP3.04"
/note="potential splice acceptor sequence"
complement(32669..32674)
/gene="MALP3.04"
/note="potential splice donor sequence, aag/gtatac"
join(36657..36743,36864..37343)
/gene="MALP3.05"
join(36657..36743,36864..37343)
/gene="MALP3.05"
/note="MALP3.05, hypothetical protein, len: 188 aa"
/codon_start=1
/product="hypothetical protein, MALP3.05"
/protein_id="CAB63560.1"
/db_xref="GI:6594248"
/translation="MRKMSGIFPKILKICISICVFECKMKISYRKDLWYSEN CFNYSIDSLAESSEKTKYKIDPINELLSLININEYKMEIYOSFMDNNLN IANVLRKNSHTENIEFSLINDSSKSPVLTKFKEGSIPLMLNNPKLLPDCO RNPCHIIIGLYIILAIYVFENFKNFEC"
36744..36749
/gene="MALP3.05"
/note="potential splice donor sequence, aag/gtatac"
36854..36863
/gene="MALP3.05"
/note="potential splice acceptor sequence"
complement(join(38049..39995,40210..40284))
/gene="garp"
complement(join(38049..39995,40210..40284))
/gene="garp"
/note="MALP3.06, garp, len: 673 aa, similarity: almost identical to GARP_PLAIF (678 aa), fasta scores: 97.68 identity in 678 aa overlap"
/codon_start=1
/product="hypothetical garp protein"
/protein_id="CAB63561.1"
/db_xref="GI:6594249"
/translation="MNVLFISYNICILFVCTLNSTKCFSGNLIKNNIINKSPDS ITGRILNTELEKNDKNSKSETLKEEDKEDVPTSDNLKRNHNNESSSDP TNILVNDKDNESVDKKKDKKREKHKKKKDKKREKHKKKKREKHKKKK DKKEENSEVMSLYTGOHKPKNATHEGSELYEAMVSINNAGGILLSPYRE OGCGGIISSVHETSDNDKDNKENSSEKEDHODDEMLTKDERKKEKMEKOE KIEKKKOEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEK HDKNEETMOODPOTSEETNNEIWPPLDPLDVTPEHKGSEKKEKKEKKEG EHKREHKEEKKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKE DKGVEITINLEDEKACEOHITVESRPSQPCKLIDEPQLLNDKSKVEKKNISIQ EQLITIGVNVPRDNRKRMKMAIEELQOKHVDKEDKESKKEVEESKEVQ EDEBEVEDDA EDDAEDDDAEDDEDEDEDEDEDEDEDEDESEKKIKRMLRNNAKI"
complement(39996..40005)
/gene="garp"
/note="potential splice acceptor sequence"
complement(40204..40209)

gene /gene="garp"
/note="potential splice donor sequence, aag/gtatac"
join(45401..46396,46562..50233)
/gene="MALP3.07"
Query Match 6.0%; Score 172.4; DB 3; Length 67970;
Best Local Similarity 46.1%; Pred. No. 1.1e-14;
Matches 787; Conservative 0; Mismatches 905; Indels 17; Gaps 6;
QY 299 atatttaacgtgattgaagcgtatccctttaaacttaataataaccagcgtc 358
Db 10033 ATATATTAATTTATTTAATAATAATATATTAATTAATTAATTAATAAAG 9974
QY 359 ttaacttaatttngcgtatgtaacgttcaacgttgcctctttttgancggt 418
Db 9973 AAATTAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9914
QY 419 gattggttatagtcttatttgaatgngcnaagtaagaataagaacaattatg 478
Db 9913 TAAGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9854
QY 479 gagaaataatatttaattgtaacataatagagaataatataatnccgtgattacg 538
Db 9853 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9794
QY 539 tatgatacgagtagaagatcttgataataatctgagaactgaccttccaaagt 598
Db 9793 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9734
QY 599 aaaaatat-ttgatataagtaagtaacataacataagaataaaaaataaatcaacaa 657
Db 9733 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9674
QY 658 atgaaaaaactgtagtgatctacccctcaagctttgaactatcttggttaccgcc 717
Db 9673 AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9614
QY 718 ctaacctcaagcttcaacaaataaaattcatatgca-----tatctatactc 772
Db 9613 TTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9554
QY 773 tttagaagtgtaacaanaatatatacaagttatataatgltttccaataaaaaagataaa 832
Db 9553 TTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9494
QY 833 aaataataaaaaataatagtagttacaataaaaaataaatatattttccaaggtca 892
Db 9493 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9434
QY 893 naaacaactaaacctaaccctaataatataacatttagttaaaccttaacgttga 952
Db 9433 ATTAATAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9374
QY 953 taaacttaaaacataaacaataaacaactaaacccaaactccaaacttaaacctta 1012
Db 9373 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9314
QY 1013 agagtttaagtggttgggttttggattatagtttagttatccaaaggtttaggt 1072
Db 9313 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9254
QY 1073 ttaaccaagatcttaggtttaggaataagtaacttaggaattgltttcaactgaagc 1132
Db 9253 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9194
QY 1133 ttcaagatcttttaaaaaatctt-----tttgtaacaactactatttattta 1186
Db 9193 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9134
QY 1187 ttttttaccttttataataaacaataataataatttaactcaactgcttcatatt 1246
Db 9133 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9074

OY 1002 ctaaaccccttaagtgcttaaatggttgattgattagattagattccaa 1061
 DB 85724 ATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85665
 OY 1062 aggttaaggttaccagaaggttaaggttaggattagattagattgatt 1121
 DB 85664 TAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85605
 OY 1122 tactgaagcgttcaagatttttaaaatatttttttgtaaacactattttt 1181
 DB 85604 AATAATTAATTTATGTTATTTATTAATTAATTAATTAATTAATTAATTA 85545
 OY 1182 attattttttaccctttatataaacaataataacttaactcactcgttc 1241
 DB 85544 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85485
 OY 1242 atatttaagtgcatgttaacattatttttggttaaaaaaatgtcaacttgaattc 1301
 DB 85484 TAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85425
 OY 1302 caatgcaaa-attattatttttcaagctaaatgaatgcaagtgatgactata 1360
 DB 85424 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85365
 OY 1361 aataatttatttctcaaatgcatattgtgcaaacatgtaataaagaactta 1420
 DB 85364 GTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85305
 OY 1421 attattatttatttttttttttttttttttttttttttttttttttttt 1480
 DB 85304 TATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85245
 OY 1481 aaacgaattgagtaataattgttcttatttttaaaagatacgaatacgaacaca 1540
 DB 85244 TTATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85185
 OY 1541 atttattgtctgtagaacctaaatcaccctcaagaagtgagcgaagaataagta 1600
 DB 85184 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85125
 OY 1601 acgttttgagggaagcctaactgagcccaagtaacataaactcttcttgatttaca 1660
 DB 85124 ATTAATTTGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 85065
 OY 1661 aatcccttaacaatttagttagagtaatacgaacaacacatgattatcatatataag 1720
 DB 85064 ATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85005
 OY 1721 tattcttaaaaaattaccagcgaataaataactcttcttatttatacaaatccctat 1780
 DB 85004 TAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 84947
 OY 1781 aaaaactatttatataactaaacattttaaaagaataaagaagaccatgatt 1840
 DB 84946 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84887
 OY 1841 acataaataatagttattcttaagtagtaataataataataataataataataata 1900
 DB 84886 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84828
 OY 1901 atttacaataataataatttttgcgttagccgataactcattactataaatacataa 1960
 DB 84829 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84770
 OY 1961 ccacatgtagatataattttatttattatataatataa 2001
 DB 84769 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 84729

RESULT 13
 AC093899 172816 bp DNA linear PRI 05-FEB-2002
 LOCUS AC093899
 DEFINITION Homo sapiens chromosome 2 clone RP11-724016, complete sequence.

ACCESSION AC093899 AC068884
 VERSION AC093899.3 GI:18497265
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 172816)
 AUTHORS Waterston,R.H.
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 172816)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (10-SEP-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 3 (bases 1 to 172816)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (05-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 COMMENT On Feb 5, 2002 this sequence version replaced gi:15625013.
 FEATURES
 source Location/Qualifiers
 1..172816
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="2"
 /clone="RP11-724016"
 BASE COUNT 53405 a 33788 c 33104 g 52519 t
 ORIGIN
 Query Match 5 9%; Score 169.6; DB 9; Length 172816;
 Best Local Similarity 46.9%; Pred. No. 2.2e-14;
 Matches 762; Conservative 0; Mismatches 840; Indels 22; Gaps 7;
 OY 382 ttatcgtgtttagtgcnccttttttttgancgcgttgattgattgattgatt 441
 DB 22715 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 22774
 OY 442 gttgcacgttaagaagaacattatataatttgagaaataataataatgatt 501
 DB 22775 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 22834
 OY 502 caataatagagaataatataatccttgctgctgctgctgctgctgctgctgctgct 561
 DB 22835 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 22894
 OY 562 tgaataatattggaacttgcttctcacaagaataaattgattgattgattgatt 621
 DB 22895 TATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 22950
 OY 622 gttacacatgagaataataaataataataataataataataataataataata 681
 DB 22951 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 23010
 OY 682 cccctcaacgttttgaaactattcttgctgctgctgctgctgctgctgctgctgct 741
 DB 23011 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 23070
 OY 742 ataaattcattatgcatattcattatcattttagaagaagaacaaataataataa 801

Db	23071	TTTACATATTAATATATTTATATTTAAATATATATGCAATATGTAATATATATTTATTTAA	23130
Oy	802	ttatactatgcttccaataaagaataaataaataaataaagtagttagtcaaa	861
Db	23131	TAAATTTATATATTAATAATATTTTATATTTATATATATTTATATTAATATATTTATAT	23190
Oy	862	aaaaaaaataatatttttaccgcgctcnaaagaactaaaccctaaccctaatat	921
Db	23191	TTAAATATATTTAAATATATATTTTATATTTATATATATTTAAATATTTAAATATATTTTAT	23250
Oy	922	taaacctttagtgaaccctaaaccttgggataaacctcctaaacatlaaacatlaaacac	981
Db	23251	ATTAAATATATATATATATTAATAACATATATTTATATTTAAATATATATATATTAATAACAT	23310
Oy	982	taaacctcctaaacctcctaaacctttaaacccttaagtgcttaaatgcttagtgc---	1038
Db	23311	ATATTTATATTTAAATATATATTTATATTAATAACATATATTTATATTTAAATATATTTTATA	23370
Oy	1039	tttaagcttagtgaattacccaaggtttaaaggttaccagaaggtttagttagttagga	1098
Db	23371	TTTATATATATATATATTAATAACATATATTTATATTTAAATATATTTATATTAATAATATA	23430
Oy	1099	ctagacttaggacttagtgcttctaacgacgcttcaagaatattttaaataatct	1158
Db	23431	TTATATATTAACATATATTTATATTTATATTAATAATATTTATATTAATAATATATTTATAT	23490
Oy	1159	tttttgtaacaactcatttctatttcttttttacccttttatataaataaactaata	1218
Db	23491	AATATATATATATTAATAATATATTTATATTTAAATATATTTATATTTAAATATATTTAT	23550
Oy	1219	taatttaactcactcgttctcatataagtgctcatgta--acattatttlltgc	1275
Db	23551	ATATTAATA-TATTTATATTTAAATATATTTATATTAATAATATTAATAATATTTATAT	23609
Oy	1276	acaaaaaatgctcctttagaatccaatccaagaacttatctatttcttcgctcaaat	1335
Db	23610	ATATTTATATTTATATATATATACATATATTTATATTTAAATATATACATATATATTTAT	23669
Oy	1336	aattgcaaaagtcattgactctataataaatttctatctcctaaagctatattgctca	1395
Db	23670	ATTTAATATATACATATATTTATATTTAAATATATACATATATATTTATATTTAAATA	23729
Oy	1396	aacatgcttaattatagaaactaatatataattcatatttcttcttaactctgtgtaa	1455
Db	23730	TAAATTTATATTTTATATATATTAATAATA-----TATTTATATTTAAATATATTA	23782
Oy	1456	aaatgccaagtaaaattatttagaaagaaagaaagaaatatttgccttctttaa	1515
Db	23783	ATATATATTTATATTTATATATATATTTATATTAATAATATACATTAATATATATTTAT	23842
Oy	1516	aagatactgaatatgaataacacaaatttattgtatagatgaaccataaattcctcta	1575
Db	23843	ATTTAATATTAATAACATATATTTATATTTATATTAATAATATATATTTATATTAATA	23902
Oy	1576	agaagtggaacgcaagaaatagctcaagctttagggaaagcctaactatggcccaagca	1635
Db	23903	TAAATATATATTTATATTTAAATATTTATATATATATTAATAATATTTATATTTAAATAT	23962
Oy	1636	tcaaaatcttcttgatcttccaataacccctacaactttagttagttagttagaccaa	1695
Db	23963	TATATATATGATATATTTTATATTTATATTTATATGATATATTAATATTTATTTAT	24019
Oy	1696	acacatgattatcatcatatagaaatcttcaaaaaattactcagcgaataatlaaacct	1755
Db	24020	ATATTTATATATTAATAATATATTTATATTTATATTTATATTAATAATATTTATATTTAT	24079
Oy	1756	tctcttatcttccaataacccctataaaaacttattatatactactaaacatttcat	1815
Db	24080	ATATTTATATTTTATATATTTTATATTAATATATATTTATTTATTTA-TATATATTTTATAT	24138
Oy	1816	taaaagaataaaggacatgagacataaaatatatgctattcttcttaagatagct	1875

Db	24139	ATAAATAATTAAATTTTATATATTATTAATTAATAAATTTAAATTTTAAACAGTTAT	24138
Oy	1876	aataattaatataaccagtcocatattatltcaaaaataaatatlttgcgtacga	1935
Db	24139	ATAAATAATTAAATTTTATTAACAGTTAATAAATAATTTATTAATTTAACTTAATAAA	24258
Oy	1936	taactattcactaaattcatcaaaacccagtagatgatatlttatataatat	1995
Db	24259	ATAATTTTAATAATTTTATACAGTTAATAAATAATTTTAAATTTTAAACAGTTAATAAAAT	24318
Oy	1996	atat 1999 	
Db	24319	ATAAT 24322	

```

RESULT 14
AE001398 14867 bp DNA linear INV 06-NOV-1998
DEFINITION Plasmodium falciparum chromosome 2, section 35 of the
ACCESSION AE001398 AE001362
VERSION AE001398.1 GI:3845197
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITILE malaria parasite P. falciparum.
Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1 (bases 1 to 14867)
Gardner,M.J., Shellen,I.H., Carucci,D.J., Cummings,L.M., Aravind,L.,
Koonin,E.V., Shellen,S., Mason,T., Yu,K., Fujii,C., Pederson,J.,
Shen,K., Jing,J., Aston,C., Lai,Z., Schwartz,D.C., Pertea,M.,
Salzberg,S., Zhou,L., Sutton,G.G., Clayton,R., White,O.,
Smith,H.O., Fraser,C.M., Hoffman,S.L. et al.
Chromosome 2 sequence of the human malaria parasite Plasmodium
falciparum
Science 282 (5391), 1126-1132 (1998)
JOURNAL
MEDLINE 99021743
ERRATUM:[[published erratum appears in Science 1998 Dec
4;282(5395):1827]]
2 (bases 1 to 14867)
Gardner,M.J.
REFERENCE
AUTHORS Direct Submission
TITILE Submitted (02-NOV-1998) The Institute for Genomic Research, 9712
JOURNAL Medical Center Drive, Rockville, MD 20814, USA
FEATURES
source
1. 14867
/origanism="Plasmodium falciparum"
/db_xref="taxon:5833"
/chromosome="2"
/complement(1570..2424)
/gene="PFB0490C"
/complement(1570..2424)
/gene="PFB0490C"
/note="predicted by glimmerM"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC71887.1"
/db_xref="GI:3845198"
/translation="MKKEKKEIMDIYSCPLDDVVDREKSKGNSLKSSTYKSDYRK
SSISIKRDSHKKSGSFRGKSGEINRKSQGFKKPYNNRLINKNNYNNKRNHNDR
DNFGKRTSGFSGRYEDNRKSGEFKKRFTISNRKSSVKSYKSGSKMGKSPFNKAPTSR
TVNKKRLNNYKTVSAPKRFNNNLISLYRKRRPALMTKREKPGVTIKSSVPRKIRKK
GIKSGISKSKTRKTSKSGSKFPLKTYLSLKIYTSINKIIPSLKEQKNFEVNLPEISL
NNATTKNN"
11241..14606
/gene="PFB0495W"
11241..14606
/gene="PFB0495W"
/note="predicted by glimmerM"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC71888.1"
/db_xref="GI:3845199"
/translation="MFAIKNTVVEGVCNIFCSYLRKLYLKSNFVTLNLTNETNNEKR

```



```

Query Match      5.9%; Score 168.4; DB 3; Length 14867;
Best Local Similarity 45.5%; Pled. No. 5,2e-14;
Matches 715; Conservative 0; Mismatches 848; Indels 9; Gaps 3;

QY 469 ttatattggagagaaataataatttaattgaatgltccaatatacagagaataattatnccct 528
DB 8106 TATTTAAATGAATAAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8165
QY 529 gatgttaccgtatgatgctgcgagtgaagaagtccttgaataatatcttgagaactgccttc 588
DB 8166 TAAATATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8225
QY 589 ctcaaaatbtaaaataatttgatcatgtaacctaaagttaacaactgaagaactaaaaaaaat 648
DB 8226 TTTTAAAAAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8285
QY 649 taaataaaatbgaanaaacatgcatgctgctcaacctcaacgcttttgaaactatcttg 708
DB 8286 TAATTTAAATTTAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8345
QY 709 gtacacccccaataacgtctcaagcttccacaacaataaaatttcattatgcatactcat 768
DB 8346 AGTCATTTTATTTATTTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8405
QY 769 atcttttgaagaagtgacaacaattctcaagtatactatgcttttccaataaaaga 828
DB 8406 TGTGTTCTTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8465
QY 829 taaaaataaaataaaataaataagtagtattccaaaaaaataaataac-----atttta 883
DB 8466 TTAATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8535
QY 884 ccagcgctcnaaaacacataaaacctaaacctaaactaaatatacttttagtgaaccccaa 943
DB 8526 TTTCCAAATTAATAATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8585
QY 944 accttggataaactttaacattbaacaacttaaaacactaaacctaaacctaaacctc 1003
DB 8586 TAAATTAATTAATTTGCTTTAA-TTCAATTAATTAATAATTAATTTTATATTAATTTTAAATA 8644
QY 1004 aaaccttaagtglttaaatgttgagtgcttttgatagatagtttagattaccagaag 1063
DB 8645 ATTATGATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8704
QY 1066 gtttaaggtttaccacaagaagttatgcttttagggagatacgtactaggattagtgttta 1123
DB 8705 ACATTAATAATTTTACACAAATAATTTATTTATTTTAATTAATTAATTTTACATAATTTTAAT 8764
QY 1124 ctgacgacttcaagaagtctttttaaaaaatttttttgtaacaactactactattttat 1183

```

[illegible]

	RESULT	15				
	LOCUS	AC093899/c				
	DEFINITION	AC093899	172816 bp	DNA	linear	PRI 05-FEB-2002Z
	ACCESSION	Homo sapiens chromosome 2 clone RP11-724C16,				
	VERSION	AC093899.AC068884				
	KEYWORDS	AC093899..3 GI:18497265				
	SOURCE	HTG.				
	ORGANISM	human.				
		Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

Mon Jul 1 08:32:46 2002

QY 1988 atatatataaa 2002
|||
Db 22903 TATTAATATAAGA 22889

Search completed: June 28, 2002, 13:44:36
Job time: 8293 sec

us-09-509-188-3.rge

QY 1981 tattatataataataaaccctaagccttaccactcgataaccataaactttt 2040
DB 1981 tattatataataataaaccctaagccttaccactcgataaccataaactttt 2040
QY 2041 cttctcgttcgtaactcaagcgttcgaaagtaaaaaaataatgaagatgtcac 2100
DB 2041 cttctcgttcgtaactcaagcgttcgaaagtaaaaaaataatgaagatgtcac 2100
QY 2101 ttgttctgtatgataccttcttctaagctgtgtacatcccaagttacagcaagcaac 2160
DB 2101 ttgttctgtatgataccttcttctaagctgtgtacatcccaagttacagcaagcaac 2160
QY 2161 tagagatcaactaactaacaagagccttctctatcgcgcaacttcttcacccca 2220
DB 2161 tagagatcaactaactaacaagagccttctctatcgcgcaacttcttcacccca 2220
QY 2221 agcaacattgacggttcggtgtgctccgaaagccttccctcgagccacttcgactaa 2280
DB 2221 agcaacattgacggttcggtgtgctccgaaagccttccctcgagccacttcgactaa 2280
QY 2281 ctccgttccatctgacacaggaagtacacagatgttgaagcaagaagagatgata 2340
DB 2281 ctccgttccatctgacacaggaagtacacagatgttgaagcaagaagagatgata 2340
QY 2341 catgtttatgatcgcgttgagacttcttccacaggaagcgttatgtgacggaat 2400
DB 2341 catgtttatgatcgcgttgagacttcttccacaggaagcgttatgtgacggaat 2400
QY 2401 gttgcgcgcgacataaagaatgaacaaagtgtgagaagacggtcttgccttcc 2460
DB 2401 gttgcgcgcgacataaagaatgaacaaagtgtgagaagacggtcttgccttcc 2460
QY 2461 atgaccccttctgacggttcgtaagctatgatactgctccacggttgttgatctact 2520
DB 2461 atgaccccttctgacggttcgtaagctatgatactgctccacggttgttgatctact 2520
QY 2521 caactctctcttcaacagagctccttcaatgctccttcttccacaggtccttccacatgctc 2580
DB 2521 caactctctcttcaacagagctccttcaatgctccttcttccacaggtccttccacatgctc 2580
QY 2581 ctccacatgctccttcaacagagctccttcaatgctccttcaatgctccttcaatgctc 2640
DB 2581 ctccacatgctccttcaacagagctccttcaatgctccttcaatgctccttcaatgctc 2640
QY 2641 ctctacatgctccttcaacagagccttcaacagccttcaacagccttcaacatgctc 2700
DB 2641 ctctacatgctccttcaacagagccttcaacagccttcaacagccttcaacatgctc 2700
QY 2701 ctctacatgctccttcaacagagccttcaacagccttcaacagccttcaacatgctc 2760
DB 2701 ctctacatgctccttcaacagagccttcaacagccttcaacagccttcaacatgctc 2760
QY 2761 tattcaagatgatagtcctaggaagcttcttcttcttcttcttcttcttcttcttctt 2820
DB 2761 tattcaagatgatagtcctaggaagcttcttcttcttcttcttcttcttcttcttctt 2820
QY 2821 tatttgaagaaacgcttcttgaatttaact 2853
DB 2821 tatttgaagaaacgcttcttgaatttaact 2853

RESULT 2

ID AAX26841 standard; cDNA; 674 BP.

AC AAX26841;

DT 22-JUN-1999 (first entry)

DE Microspore-specific plant promoter cDNA clone M3.21.

KW Microspore-specific plant promoter; Brassica napus cv. Brutor;

KW gametophytic male sterility; inducible fertility; ss.
XX Brassica napus.
OS
XX
PN FR2768745-A1.
XX
PD 26-MAR-1999.
XX
PF 23-SEP-1997; 97FR-0011812.
XX
PR 23-SEP-1997; 97FR-0011812.
XX
PA (INRG) INRA INST NAT, RECH AGRONOMIQUE.
PI Drouaud J, Fourgoux A, Guerche P, Pelletier G.
XX
DR WPI; 1999-217496/19.
XX
PT Microspore-specific plant promoter from Brassica napus - and vectors
PT for producing plants exhibiting gametophytic male sterility with
PT inducible fertility
XX
PS Example 1; Page 14; 32pp; French.
XX
CC The present sequence represents a clone of a microspore-specific plant
CC promoter from Brassica napus cv. Brutor. The promoter is used in a
CC method for producing plants exhibiting gametophytic male sterility
CC with inducible fertility.
XX
SQ Sequence 674 BP; 155 A; 181 C; 124 G; 214 T; 0 other;

Query Match 16.2%; Score 462; DB 20; Length 674;

Best Local Similarity 81.9%; Pred. NO. 4.2e-54;

Matches 614; Conservative 0; Mismatches 35; Indels 101; Gaps 3;

QY 2105 tcttctatgatcctcttcttcaagctgtgtacacacaaagtacagcaagaactaga 2164
DB 1 tcttctatgatcttcttcaagctgtgtacacacaaagtacagcaagaactaga 60
QY 2165 gtcatcaactaacaagaagctcctctatcgcggaacttccctgcttaccaccaaga 2224
DB 61 gtcatcaactaacaagaagctcctctatcgcggaacttccctgcttaccaccaaga 120
QY 2225 acattgacggttcggtgctccgaaagccttccctcgagcgaacttcgacttaactc 2284
DB 121 acattgacggttcggtgctccgaaagccttccctcgagcgaacttcgacttaactc 180
QY 2285 gtccatctgccaacaggaatccacagatgttgaacagcaagaagaggttaggtacatg 2344
DB 181 gtccatctgccaacaggaatccacagatgttgaacagcaagaagaggttaggtacatg 240
QY 2345 tttaatgatatgctgagacttcttccacaggaagcgcgcatgtgacggaatgtg 2404
DB 241 tttaatgatatgctgagacttcttccacaggaagcgcgcatgtgacggaatgtg 300
QY 2405 cgcgcgcatcaagaagatgacaagaagatgtggaagaacgcttcttgaatcttccatga 2464
DB 301 cgcgcgcatcaagaagatgacaagaagatgtggaagaacgcttcttgaatcttccatga 360
QY 2465 cccctctctgacggtatgtcaagctacatgtctccacggtgtgtgactactaac 2524
DB 361 cccctctctgacggtatgtcaagctacatgtctccacggtgtgtgactactaac 420
QY 2525 tctctcttcaagcgtccttcaatgctccttcttcaagcgtccttcaatgctcctc 2584
DB 421 tctctctc----- 429
QY 2585 acatgctccttcaagcgtccttcaatgctccttcaatgctccttcaatgctcctc 2644
DB 430 -----acatgctcctc 441
QY 2645 acatgctccttcaagcgtccttcaagcgtccttcaagcgtccttcaatgctcctc 2704

Db 442 acgggccccttaccatgctcctcaccaggtccttaccgctccctcaccaggtcctt 501
QY 2705 acgcccccttcgaagctcctcaccggtcagta-cttgatctttgagaattat 2763
Db 502 acgcccccttcaccagctcctccaccggtcaggtatcttgatctttgagaattat 561
QY 2764 tcaagattgagtctcctcagggagtttcttctctgttttaaatcttgattat 2823
Db 562 tcaagtgatgctcctcagggagtttag-gtttctctgttttaaatcttgattat 620
QY 2824 ttgagaaaacggtcttgatttactt 2853
Db 621 ttgagaaaacggtcttgatttactt 650

RESULT 3
AAZ6840
ID AAZ6840 standard; cDNA; 497 BP.
XX
AC AAZ6840;
XX
DT 22-JUN-1999 (first entry)
XX
DE Microspore-specific plant promoter cDNA clone M3.
XX
KM Microspore-specific plant promoter; Brassica napus cv. Brutor;
KM gametophytic male sterility; inducible fertility; ss.
XX
OS Brassica napus.
XX
PN FR2768745-A1.
XX
PD 26-MAR-1999.
XX
PE 23-SEP-1997; 97FR-0011812.
XX
PR 23-SEP-1997; 97FR-0011812.
XX
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
XX
P1 Drouaud J, Fourgoux A, Guerche P, Pelletier G;
XX
DR WPI; 1999-217496/19.
XX
PT Microspore-specific plant promoter, from Brassica napus - and vectors
PT for producing plants exhibiting gametophytic male sterility with
PT inducible fertility
XX
PS Example 1; Page 13-14; 32pp; French.
XX
CC The present sequence represents a clone of a microspore-specific plant
CC promoter from Brassica napus cv. Brutor. The promoter is used in a
CC method for producing plants exhibiting gametophytic male sterility
CC with inducible fertility.
XX
SQ Sequence 497 BP; 107 A; 130 C; 70 G; 190 T; 0 other.

Query Match 13.8%; Score 393.2; DB 20; Length 497;
Best Local Similarity 98.0%; Pred.No. 7.7e-45;
Matches 398; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2448 ttggatcttcacatgacccctcttgaacggtatgtaagctacatgctccacgtt 2507
Db 1 ttggatcttcacatgacccctcttgaacggtatgtaagctacatgctccacgtt 60
QY 2508 gtggatcttcacatgacccctcttgaacggtatgtaagctacatgctccacgtt 2567
Db 61 gtggatcttcacatgacccctcttgaacggtatgtaagctacatgctccacgtt 120
QY 2568 ccttacaatgctcctcaccaggtcctccttacaaggtccttacaatgctccttacaatgct 2627

Db 121 ccttacaatgctcctcaccaggtcctcctcaccaggtccttacaatgctccttacaatgct 180
QY 2628 ccttacaatgctccttaccaggtcctcctcaccaggtccttacaatgctccttacaatgct 2687
Db 181 ccttacaatgctccttaccaggtcctcctcaccaggtccttacaatgctccttacaatgct 240
QY 2688 ccttacaatgctccttaccaggtcctcctcaccaggtccttacaatgctccttacaatgct 2747
Db 241 ccttacaatgctccttaccaggtcctcctcaccaggtccttacaatgctccttacaatgct 300
QY 2748 tattgataagattatcaccaggtatgctcctcaccaggtatgatttttttttttttttttt 2807
Db 301 tattgataagattatcaccaggtatgctcctcaccaggtatgatttttttttttttttttt 360
QY 2808 aaatttgggttattttagaanaacggtccttgatttactt 2853
Db 361 aaatttgggttattttagaanaacggtccttgatttactt 406

RESULT 4
AAC34997
ID AAC34997 standard; DNA; 730 BP.
XX
AC AAC34997;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 8641.
XX
KM Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PF 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 28-APR-1999; 99US-0130510.
PR 30-APR-1999; 99US-0130891.
PR 03-MAY-1999; 99US-0131449.
PR 16-MAY-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132486.
PR 08-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154479.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

PR	13-OCT-1999;	99US-0159294;
PR	13-OCT-1999;	99US-0159295;
PR	14-OCT-1999;	99US-0159329;
PR	14-OCT-1999;	99US-0159330;
PR	14-OCT-1999;	99US-0159331;
PR	14-OCT-1999;	99US-0159637;
PR	14-OCT-1999;	99US-0159638;
PR	18-OCT-1999;	99US-0159584;
PR	21-OCT-1999;	99US-0160741;
PR	21-OCT-1999;	99US-0160767;
PR	21-OCT-1999;	99US-0160768;
PR	21-OCT-1999;	99US-0160770;
PR	21-OCT-1999;	99US-0160814;
PR	22-OCT-1999;	99US-0160980;
PR	22-OCT-1999;	99US-0160981;
PR	22-OCT-1999;	99US-0160989;
PR	25-OCT-1999;	99US-0161404;
PR	25-OCT-1999;	99US-0161405;
PR	25-OCT-1999;	99US-0161406;
PR	26-OCT-1999;	99US-0161359;
PR	26-OCT-1999;	99US-0161360;
PR	26-OCT-1999;	99US-0161361;
PR	28-OCT-1999;	99US-0161920;
PR	28-OCT-1999;	99US-0161992;
PR	28-OCT-1999;	99US-0161993;
PR	29-OCT-1999;	99US-0162142.

Query Match 6.5% Score 184.4; DB 21; Length 461;
Best Local Similarity 67.9%; Pred No: 1e-16;
Matches 311; Conservative 0; Mismatches 126; Indels 21; Gaps

OY	2104	tctctgtatgatcctcttcttaagctgtygtacataccaagttaacgaacaagaactg	2163
Db	1	tgttgcatataactcttctctaagctgtygtacatctcaagtcacgaacgaanaaatg	60
OY	2164	agatcataac-----aaaccaagagctcttctatcggcga-----ctaacctc	2208
Db	61	attatcttcatttagctacgaagagctcgatgtygtacattacaactcttttaccttc	120
OY	2209	gcttcaccaccaagaacattggccgttcgtygtcccggaagaagcttcacc-----tg	2265
Db	121	acctttacccttaaggcacatcttgccatctcccccaccccggaagagctctccaccaatcc	180
OY	2263	caggccacttcgcgactaacctcgcttcacatcgcgaaggaatcacagaatgcttgaacg	2322
Db	181	cagcgcgcttcaccacgcgattccatccatccaccaccgcgttgaaccaaatgcttgcgcg	240
OY	2323	acaaagaaggagtgatgtagatggtttaatgatatgcgtcagacctcttcaccaggaag	2388
Db	241	attcgaagagtgataaacaactglttgcgagatgcgttaaagcttcttcaccgcgaac	300
OY	2383	ccgctatgttatcggaatgtytgcgccgcgataagaagaatgataacaagaatttggaga	2444
Db	301	cagctatgtgatgttgatgttgttcgcgccctccatccaagaagaatgatgaaatttgaata	360
OY	2443	ccgctcttgatctacacagccccctttgcacgcggcatgtcaagactacatatgtcca	2502
Db	361	ctgctctttggagcctaccacaacaccttctcttgcagctgtygttaagcttcattgctcta	420
OY	2503	ccgctgttgatctactacacccctccctccctcaagaagct	2540
Db	421	ccaagctaatccatccatccgtcgtctccatccatccagct	458

RESULT 6.
AAx26842/c
ID AAx26842 standard; cDNA; 2853 BP.
XX AAX26842;
XX AC
XX DT 22-JUN-1999 (first entry)

	Query Match	6.1%	Score 174:8:	DB 20:	Length 2853;
	Best Local Similarity	68.6%;	Pred: No. 1.9e-15;		
	Matches 310;	Conservative 0;	Mismatches 114;	Indels 28;	Gaps
OY	793 atatacaagtcatactatagtgtttccaataaaagaataaaaataataaaataatag	852			
Db	1230 AGTATTAAATTAATTAATGTTTATAATATAAAGGTAAAAAATTAATTAATTAATGAG	1171			
OY	853 tagtacaacaaaaa-----aatlaattttaccagcgt-----canaaac	898			
Db	1170 TGTGTACAAAAAAAATATTTTTTAAAAAATACTTTGAACGTCTGCTAGTAACAACATAAT	1111			
OY	899 actaaacctaaaccttaatatltiaacttttagtguaaacccctaaaccttgataaac	958			
Db	1110 CCGAAGTCATATTCCTTAACCATTAACCTCTGGGTAACCTTAAACCCTTTTGCAATAAATC	1051			
OY	959 ttaaaccattaaacattaaacacacaaacctaaacctaaacctcaaacaccttagtgt	1018			
Db	1050 CTAAACTATATAATTCAAAAAACCTAAACATTTTAAACACTTAAAGGTTTAGAGTTAGATT	991.			
OY	1019 taaggttagtggttttgattatagtttaggatttatccaaaggttlaagtttaacc	1078			
Db	990 TAGGGTTTAGTGTTTAATGTTTAATGTTTAATGTTTAAGATTTTCCAAAGCTTTAGGGTTTACCT	931			
OY	1079 aaggattataggtttaggaattatactacttagatttagtgytttacttaggaacgttcaa	1138			
Db	930 AAAAGTTTAATTAATTTAGSGTTTAGGTTTAG-----TGTTTNTTATACGCTGGTAAAA	879			
OY	1139 gtaattttaaaaaaatcttttttggtaacgaactacttttlatitatttttacct	1198			
Db	878 ATATT-----AATTTTTTTTTTTTGTGAACACATATATTTTATTAATTTTATTCCTT	825			
OY	1199 ttatataaaaaacataataataact	1230			
Db	824 TTATTTGAAAAACATATATTAACCTTGATTAAT	793			

ID	ABL32315/c
XX	ABL32315 standard; DNA; 6641 BP;
AC	ABL32315;
AG	
AT	26-MAR-2002 (first entry)
CC	
DE	Human immune system associated gene SEQ ID NO: 288.
GG	
HH	Human; immune system disease; cytosine methylation; antiasthmatic;
KK	antiarteriosclerotic; antihaemiac; cytostatic; neoptropic;
KW	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW	antiinflammatory; antiarthritic; antidiabetic; antipsoriatic;
KW	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW	neutrophilomatosis; rheumatoid arthritis; psoriasis; bowel disease;
MM	gene; ds.
OS	Homo sapiens.
PN	-MO200200928-A2.
XX	
PD	03-JAN-2002.
XX	
PF	02-JUL-2001; 2001WO-EP07537.
XX	
PR	30-JUN-2000; 2000DE-1032529.
XX	
PR	01-SEP-2000; 2000DE-1043826.
XX	
PA	(EPIG-) EPIGENOMICS AG.
PI	Olek A, Plepenbrock C, Berlin K;
XX	
DR	WPT; 2002-130909/17.
XX	
PT	Nucleic acid comprising fragment of chemically modified gene, useful
XX	for diagnosis and treatment of diseases associated with abnormal
XX	cytosine methylation
PS	
XX	Claim 1; SEQ ID NO 288; 32pp + Sequence Listing; German.
CC	The present invention provides a number of human immune system associated
CC	genes which are modified by the methylation of cytosines. The sequences
CC	can be used in the diagnosis and treatment of immune system disorders,
CC	including eye diseases such as retinopathy, neovascular glaucoma and
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC	leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC	diseases. The present sequence is a gene of the invention.
XX	
SQ	Sequence 6641 BP; 2495 A; 22 C; 984 G; 3140 T; 0 other;
Query Match	4.3%; Score 123.6; DB 24; Length 6641;
Best Local Similarity	46.2%; Pred. No. 1.4e-08;
Matches 604; Conservative	0; Mismatches 680; Indels 23; Gaps 5
Y	737 aaacaataaaatttcacatcattgcatcttgcatacttttagagaagtgaaaaaatatta 796
Dd	3626 AAACCTAAATTTGTACTAATCTTATTCTTACTCGGAATTAATCATATATCTCA 3567
Y	797 tcaagtatatctgttttccaataaagaataaaaataaataaataaataagtagt 856
Dd	3566 TTTTCCTTTTCACMAATTCACCTCAATTTCTTAATAATACTATCTATATATTTCCOTT 3507
Y	857 tacaaaaaaaaaaaaaataaataattttaccagcgctcnaaaaacactaaacctaaacctta 916
Dd	3506 TCCATATTAATAAAAAAATATATATACATTTTA--CATAAATTTTTATNAAAATTCATPAA 3449
Y	917 aataataaaccttttaggaacccctaacccttggataaactttaaacattaaactiaa 976
Dd	3448 ACACTTAACCTCATTAATAATTCACAATATATATATATATATATATATATATATATAT 3390

QY 977 aacacaaacccta---aaccctaaactaaacccttaagtggttaagtggt 1032
 DB 3388 CTTATATAATATAATATAATATAATATAATATAATATAATATAATATA 3329
 QY 1033 ttgtgattatagtttagattatccaaagtttaagtttccaaagattatggt 1092
 DB 3328 CTAATATATAATATAATATAATATAATATAATATAATATAATATAATATA 3269
 QY 1093 tagggatttagttaggttaggttttagttagttagttagttagttagttagt 1152
 DB 3268 AAAATCCAAATATATCTATTTATATAATATAATATAATATAATATAATATA 3209
 QY 1153 atattttttttagttagttagttagttagttagttagttagttagttagttagt 1211
 DB 3208 TATATCTATATAATATAATATAATATAATATAATATAATATAATATAATATA 3149
 QY 1212 cataataataataataataataataataataataataataataataataata 1271
 DB 3148 TATATATCTATATAATATAATATAATATAATATAATATAATATAATATAATATA 3089
 QY 1272 tttttagttagttagttagttagttagttagttagttagttagttagttagttagt 1331
 DB 3088 TATATAATATAATATAATATAATATAATATAATATAATATAATATAATATAATATA 3029
 QY 1332 aatttaattgcaagtgatgatttatttatttatttatttatttatttatttattt 1391
 DB 3028 TATATAATATAATATAATATAATATAATATAATATAATATAATATAATATAATATA 2969
 QY 1392 gtcaacacgtgtaatttaataataataataataataataataataataata 1451
 DB 2968 TTTAAATCAACAAAAATATCAATATAATATAATATAATATAATATAATATA 2909
 QY 1452 gtcaaaagtgcgaagtaatttattttagttagttagttagttagttagttagt 1511
 DB 2908 TTTATATCTCTCATATATCTT-----TCTATATAATATAATATAATATAATATA 2864
 QY 1512 ttaaaagatatgcaatataataataataataataataataataataataataata 1571
 DB 2863 AACATATTAACCTCAAAATACCTAAATAATATAATATAATATAATATAATATA 2804
 QY 1572 cctaagaagtgcaagcaagaataagtcgaagtttggtgggaagtcgaactatg 1631
 DB 2803 TCATATATCTCAATATATCTTTTAAATAATATAATATAATATAATATAATATA 2744
 QY 1632 gtcaataaaacttttctgtatttcaaaacttcaaaacttcaaaacttcaaaactt 1691
 DB 2743 TAAATATACTAAATATAATATAATATAATATAATATAATATAATATAATATA 2684
 QY 1692 ccaaacacatgattatcatatattgaatttcaaaacttcaaaacttcaaaactt 1751
 DB 2683 CCAATATCAACCTCAATATCTCTTATATAATATAATATAATATAATATAATATA 2624
 QY 1752 atcttcttatttatttcaaaacttcaaaacttcaaaacttcaaaacttcaaaactt 1811
 DB 2623 TATATATATAATATAATATAATATAATATAATATAATATAATATAATATAATATA 2564
 QY 1812 taattaaagaataaaggagacatgatacaataaataatagttatttcttaagatag 1871
 DB 2563 TATATATATAATATAATATAATATAATATAATATAATATAATATAATATAATATA 2504
 QY 1872 tgaataatataataataataataataataataataataataataataataataata 1931
 DB 2503 TATATATATAATATAATATAATATAATATAATATAATATAATATAATATAATATA 2444
 QY 1932 cgaataacatataataataataataataataataataataataataataataata 1990
 DB 2443 TACGTATATAATATAATATAATATAATATAATATAATATAATATAATATAATATA 2384
 QY 1991 tatatatataaaccttaacgcttcaacgcttcaacgcttcaacgcttcaacgct 2037
 DB 2383 TATATATATATCTCTCTATATAATATAATATAATATAATATAATATAATATA 2337

RESULT 8
 ABL34155/c
 ID ABL34155 standard; DNA: 15548 BP.
 XX
 AC ABL34155;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human immune system associated gene SPO ID NO: 2128.
 XX
 KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antihaemic; cytosolic; neotropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;
 KW antineumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antineurotic; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 02-JUL-2001; 2001WO-EP07537.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIC-) EPIDENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI: 2002-130909/17.
 XX
 PT Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation
 XX
 PS Claim 1; SEQ ID NO 2128; 32pp + Sequence Listing; German.
 XX
 SS The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 SQ Sequence 15548 BP; 4209 A; 247 C; 2903 G; 8189 T; 0 other;
 Query Match 4.3%; Score 122.4; DB 24; Length 15548;
 Best Local Similarity 44.7%; Pred. No. 2e-08;
 Matches 690; Conservative 0; Mismatches 833; Indels 19; Gaps 5;
 QY 467 aattatatttggagaataataataataataataataataataataataataataata 526
 DB 13157 AAAATATATATATATAATATAATATAATATAATATAATATAATATAATATA 13098
 QY 527 ttgatgt-tactgtatgagatgagagatgagatgagatgagatgagatgagatgagat 585
 DB 13097 TATATATATATAATATAATATAATATAATATAATATAATATAATATAATATA 13038
 QY 586 ttctcctcaaaagtaaaatatttgaatgaatgaatgaatgaatgaatgaatgaatga 645
 DB 13037 TATATATAAATAATATAATATAATATAATATAATATAATATAATATAATATA 12978
 QY 646 aattaaatcaaaatagaaaatagatgagatgagatgagatgagatgagatgagat 705
 DB 12977 AAATATATATATAAATAATATAATATAATATAATATAATATAATATAATATA 12918

[illegible][illegible]

DE	Human immune system associated gene SEQ ID NO: 490.
XX	
DT	26-MAR-2002 (first entry)
XX	
AC	ABLJ32517;
XX	
ID	ABLJ32517 standard; DNA; 5930 BP.
OS	Homo sapiens.
XX	
PN	MO200200928-A2.
XX	
PD	03-JAN-2002.
XX	
PE	02-JUL-2001; 2001WO-EP07537.
XX	
PR	30-JUN-2000; 2000DE-1032529.
XX	
PP	01-SEP-2000; 2000DE-1043826.
PA	(EPIG-) EPIGENOMICS AG.
PI	Olek A, Piepenbrock C, Berlin K;
DR	WPI; 2002-130909/17.
PT	Nucleic acid comprising fragment of chemically modified gene, useful for diagnosis and treatment of diseases associated with abnormal cytosine methylation
PS	Claim 1; SEQ ID NO 490; 32pp + Sequence Listing; German.
CC	The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel diseases. The present sequence is a gene of the invention.

Seq	Sequence	5930 BP;	2122 A;	24 C;	912 G;	2872 T;	0 other;
Query Match		4.2%	Score 121;	DB 24;	Length 5930;		
Best Local Similarity		45.8%	Pred. No. 3.2e-08;				


```
OY 1062 aggttaaggttaccagaagttttaggttaggattagcttaggttaggttt 1121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2972 TAAATAAATTAATTTCTCCCAACATCAAGTAACATTAACATCTCTT 2913
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1122 tactgaagcgttcaagaattttttaaaatatttttttgtaacactactattt 1181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2912 CACTTAACACCTTCACATCAATACCTCTCTCTTAAACCTTAA 2853
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1182 attatttttttacccttta-tattaacataataataatttaactccactggtt 1240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2852 AACCATTTTAATTAATTTCTACCTTTTAAACAAAATTTAAACCATTAACCTTCTA 2793
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1241 catattgaagtgcatgttaacattatttttgtaacaaaatgtcacttagaatt 1300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2792 TCAATATACATGACCATTAATTAATAAAATTAATCAAAATTAACCATTAACAA 2733
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1301 ccaatgcaaaattatttcttttcagctaaatattgcaaatgtcattgacttata 1360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2732 AAAACCTCACCTCTAATTAATAAAATTTAAATAAAACCAAAAAAAATTA 2673
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1361 aataattttattatctcaaatgctatattgcaacatglttaatttaagaactta 1420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2672 ACTAATTTCCATCAAAAAAAACATTAATTAATTAATTAATTAATTAATTA 2613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1421 attatttttattatttttttctta----atctglttaaaatgtcagaatattat 1476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2612 AATTAATAAAACAAAAACATTCCTAATCCATTAATTAATTAATTAATTAATTA 2553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1477 tgaagaagaattgtagtaatttcttcaattttttaaagaatcgaatgaaataa 1536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2552 CTAAATTAATTAATAAATTTAATTAATTAATTAATTAATTAATTAATTA 2493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1537 cacaattttattgtagtaagaacctaaatctccttaagaagtggaagcaagaataa 1596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2492 CATCAACAAATTAATTAATTTACATTAACCTAATAAAACAAATTAATTAATA 2433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1597 gtaacggttttggaagaacttaactggaagcaagtcacaaactcttctgtattta 1656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2432 TCCTACTACTATATATGAGAACTCTACCTTATCTTAATAAACAAATTAATTTTA 2373
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1657 tcaaaacttacaacttagtagtgaagtaagaccacaacatgattatcatcatt 1716
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2372 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2313
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1717 agaattttttaaataacttagcgaataataaactttcttattatcaaaatcc 1776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2312 TTTAAATCTTTTAAATAATCAAAATTAATTAATTAATTAATTAATTAATTA 2253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1777 ttataaaacttattatataatactaaacaaatlttaataaagaataaagagcat 1836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2252 AATTCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1837 ggaatataaaatataatgtatttcttaagatgataataataataatccagctcc 1896
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2192 AATTTAAAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTA 2137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1897 atattatttcaaaataataatatttctgtagtcgataatcatcattactaataatcat 1956
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2136 AATTAATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2077
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1957 aaaaacacatgtagatgtattattatataataataataaaccctaaagccttacc 2016
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2076 AATACCGCATCTTAATAAATTAATTAATTAATTAATTAATTAATTAATTA 2017
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2017 caactgataacatacaaaccttctctcggttcgtaactaagcctcgaaagataa 2076
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2016 ATTAACCTTAATCCAAAAACATTTTAAATATCTTAATTAATTAATTAATTA 1957
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2077 aaaaaacaatgaa 2089
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1956 AAAAAAAATTAATA 1944
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 12
ABL33307/c
ID ABL33307 standard; DNA; 6175 BP.
XX
AC ABL33307;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SPO ID NO: 1280.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antihaemic; cytosolic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritis; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001MO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation.
XX
PS Claim 1: SEQ ID NO 1280; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 6175 BP; 1506 A; 154 C; 1357 G; 3158 T; 0 other;

Query Match 4.2%; Score 119; DB 24; Length 6175;
Best Local Similarity 45.9%; Pred. No. 5.9e-08;
Matches 594; Conservative 0; Mismatches 686; Indels 14; Gaps 5;

OY 816 tcaataaanaagaataaataaataaataaataaataaataaataaataa 875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6083 TCTCAAAAAACAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTA 6024
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 876 tatltttaccaggttcanaaacaacataaaccctaaactaataatatttagtga 935
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6023 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5964
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 936 aaccctaaaccttggataaatcttaaacataaacataaacataaacctaaatcc 995
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5963 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 996 taaacttaaaccttaagtttaaatgttttagcttttttagttaagtttaggatt 1055
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


Db 15004 AAAATTTAAAAATAAAAAAATACGTCATCAAAATCTAATATCTCCAAACATATTATTT 14945
 Qy 874 aatattttaccagctcnaaaacctaaccctaaccctaaccctaaccctaaccctaacc 933
 Db 14944 AATATCAAAAAATTAATTAATTCATTTCAAAACATTAATTAATTAATTAATTAATTA 14885
 Qy 934 taacacctaaaccttctgtaaatcttaaaccttaaaccttaaaccttaaaccttaaac 993
 Db 14884 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 14825
 Qy 994 cctaaaccttaaaccttaaaccttaaaccttaaaccttaaaccttaaaccttaaac 1053
 Db 14824 ACCTTAAACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 14765
 Qy 1054 ttatccaaaggtttaaggtttaccccaaggtttaaggtttaaggtttaaggtttaag 1113
 Db 14764 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 14705
 Qy 1114 tagtctttactgaagcgttcaaggtttttaaataatcttttctgaacacta 1173
 Db 14704 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 14645
 Qy 1174 ctaattttcttacttttttacccttttataaataacataataatcttaactca 1233
 Db 14644 TTAATCTTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 14585
 Qy 1234 tctgtttctatgaagtgctatgttaactatcttttctgaacaaataatgttaact 1293
 Db 14584 TATATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 14525
 Qy 1294 tagaattccaatgaacaaatcttaattcttcaagctaaataatgtcaagtgctatga 1353
 Db 14524 TTAATACCTTCAACCAAAATTTTCAATTAATTAATTAATTAATTAATTAATTAAT 14465
 Qy 1354 tcttataataatcttattatctatcctaagctatgttcaagctatgttaactatga 1413
 Db 14464 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 14410
 Qy 1414 aaacttaattatcttactatcttcttcttcttcttcttcttcttcttcttcttct 1473
 Db 14409 TATATTTCTTTAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 14350
 Qy 1474 talttgaagaagtaagtaatttcttcttcttcttcttcttcttcttcttcttctt 1533
 Db 14349 ACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 14290
 Qy 1534 taacacaaattctatgtatgaagaaacctaaccctaagaagtgaaagcgaaga 1593
 Db 14289 CATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 14230
 Qy 1594 taagtcagaagtttggggaagcttaactatgcccgaagctcaacatcttcttctg 1653
 Db 14229 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 14170
 Qy 1654 ttatcaaaatccttacaatcttcttcttcttcttcttcttcttcttcttcttctt 1710
 Db 14169 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 14110
 Qy 1711 catattagaatcttcaaaaaatctactagcgaataatcaacatcttcttcttctat 1770
 Db 14109 AAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 14050
 Qy 1771 aaactcttaaaaaatcttcttcttcttcttcttcttcttcttcttcttcttct 1830
 Db 14049 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 13990
 Qy 1831 gaccatgatacataaaatataatgtattcttaagatgataatataatataatata 1890
 Db 13989 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 13930
 Qy 1891 cagtcataatattatcaaaaaataataatttctgtagtcgataatcaatctactata 1950

Db 13929 AAAATTTAAATCATTAACAAATTAATTAATTAATTAATTAATTAATTAATTA 13870
 Qy 1951 attcataaaaccacagtgatgatattattatataatataatataatataatata 2010
 Db 13869 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 13810
 Qy 2011 ccttaccactcgaataacacataacatcttcttcttcttcttcttcttcttct 2070
 Db 13809 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 13750
 Qy 2071 aagtaaaaaaaacataatgaagaat 2093
 Db 13749 AAACCAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 13727

RESULT 14
 ABL32155/C
 ID ABL32155 standard; DNA; 6681 BP.

AC ABL32155;
 DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 128.

KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antihaemic; cytosine; noctropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antiparasitic;
 KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.

OS Homo sapiens.

PN WO200200928-A2.

PD 03-JAN-2002.

PF 02-JUL-2001; 2001WO-EP07537.

PR 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

PA (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K.

DR WPI: 2002-130909/17.

PT Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal

PT cytosine methylation

PS Claim 1; SEQ ID NO 128; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.

CC Sequence 6681 BP; 2106 A; 52 C; 1170 G; 3353 T; 0 other;

Query Match 4.2%; Score 118.6; DB 24; Length 6681;
 Best Local Similarity 45.6%; Pred. No. 6; 6e-08;
 Matches 745; Conservative 0; Mismatches 856; Indels 34; Gaps 8;

OY 456 aatgagaagaacattatattatgtggagaaatabaatlaattatgltcaatataagaga 515
 Db 3267 ATATTTTAAAAAAATTTTAAATTAATTAACAATTTATTTAAACAAAAAACTAATATACA 3206
 OY 516 aatattatnocctgagcttaccgtgatgtgagtgagagctcttgataatttga 575
 Db 3207 TCCACTACAAATTTAAAAAACAATACAAATAAAAATTTTACACATTTAAATATCTTCACA 3148
 OY 576 gaactgctcttccaaagtaaatatttgatagtbaacttaagttaacacatgaa 635
 Db 3147 ATBACAAAAAAATTCAAAAACAAAACCTTTTAAATATATACAACAATTTTAAAAAATA 3088
 OY 636 attaaaaaaaatlaaatcaaatagaaaaaciga--tagtgctaaccttcaagct 692
 Db 3087 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3028
 OY 693 ttggaactattcttggttcaaccccttaacacctgaagttcaacaaacataatltca 752
 Db 3027 TTTAAACGCCAAAAATTTATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2968
 OY 753 ttatgcatactcatctcttttgaaagtgaaacaaatatttcaagtatacttct 812
 Db 2967 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2908
 OY 813 ttccaataaagaataaaaaataaataaataaagtagtagtctacaaaaaataaata 872
 Db 2907 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2848
 OY 873 taattatttccagcgtcamaaaacctaataacctaataacctaataatlaaacctttag 932
 Db 2847 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2794
 OY 933 gtaaacctaaaccttgagataaactttaaactttaaactttaaactttaaactttaa 992
 Db 2795 -TAAATATATATATTTCTCTACCCCTTAAACAAACATTTAAATTTTTCATTTAAAA 2737
 OY 993 tcccaactctaaaccttaagtggttaaatgttggtgttttgatttagttagga 1052
 Db 2736 ACCCAAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2677
 OY 1053 ttatccaaaggtttaaggtttaccaaaggttattggtttaggaattagcttaggat 1112
 Db 2676 AAAAAAAGAAAAACCAACAATACTACTATTAATTAATCATTT--ACTAATTTATTCGACTCA 2619
 OY 1113 ttagtgttactgagcagagcttcaaggtttttttaaataatttttttgaacact 1172
 Db 2618 AAACGCTTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2561
 OY 1173 accttattattatttttttaccttttataataaacataataatlaattactacc 1232
 Db 2560 TCTCTACTTTTAAATTAATTAATTAATTAATTAATTTTATTAATTAATTAATTAATTAAT 2501
 OY 1233 atctggttcatattagtgctatgtaacattatttttgttccaaaaaatgtcaact 1292
 Db 2500 ATATACTATCCATTAATTAATTAATTTCTTAACCAAAAGCTAAATTAATTAATTAATTAATTAAT 2441
 OY 1293 ttagaattccaatgcaaaattatattatttttcaagct---aaattatgcaagagtc 1348
 Db 2440 CTTAACTTCAAAATTAACATCTAATTAATTTTAACATTTTAAATTAATTAATTAATTAATTAAT 2381
 OY 1349 attgacttataataattattattatttctcaatcgttgaacaaatgtaacagtga 1408
 Db 2380 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2321
 OY 1409 aatagaacattatatttcaattatttttcttaacgttgaacaaatgtaacagtga 1468
 Db 2320 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2261
 OY 1469 aattatttggaaacgaattgagtaaatatttgcattttttaaagaatcagata 1528
 Db 2260 AATCTTTTAAAAAATCTTAATTTTACAAACCAATCTAAATTAATTTCTTAAAAA-----AATTT 2207
 OY 1529 tgaataacacaaatttattgtagtgaagcctaaataatcattccttaagaagtgaaagc 1586

XX WPI; 2002-130909/17.
DR
SU

PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation

PS Claim 1; SEQ ID NO 2128; 32pp + Sequence Listing; German
vz

CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.

SQ Sequence 15548 BP; 4209 A; 247 C; 2903 G; 8189 T; 0 other;

Query Match	4.28	Score 118.6	DB 24	Length 15548
Best Local Similarity	44.28	Pred. No. 6.4e+08		
Matches 748	Conservative 0	Mismatches 928	Indels 18	Gaps 6

OY	303	taaacgagagcttgaagttatcctcttttaactttaataatbaacagagctttac	362
Db	11616	ttctcttttatatttcttctttatattatcttttaatttataatttattttctt	11675
OY	363	ttatbaatttngcatnagtatacaggttttagcnccttttttttngnccgttgat	422
Db	11676	atatcttatcttattctttttttattctttatcttattctttattttatttttt	11735
OY	423	ggttaagctatttgaatgtnccnagtaagaagaacaatttatatttgsga	482
Db	11736	tatatcttatcatatatttatttttatatttttttatattttattttatttt	11795
OY	483	aaataaatttaatatgttcaatatatagagaanaattatncccttgatgtaagtg	542
Db	11796	atttttatattttatttttaatttttttttttatattttattttatttttttt	11855
OY	543	gattgagatagaagatcttttgaaataatttggagaacttgctctttccaaaagttaa	602
Db	11856	tttatcttttatatttatttatttttttttttatattttattttatttttttttt	11915
OY	603	tattgatactgaacttaagttaacacatggaanaattaaanaaaatlaaacnaatbga	662
Db	11916	ttttatttttaattttatttttttaattgttattttatcatggtatataattttttt	11975
OY	663	aaaaactgaatgtactaccctcaacggttttgaacttattcttggttcaacccttaa	722
Db	11976	atagttatataatttattttttttatagttatataattttttttt - tttatttttt	12033
OY	723	cctcgaagttcccaacaataaatttcattgcatctgcatctatcccttttggagaagt	782
Db	12034	tttatagttattttttatcatgtaattttttatcatgtaatttttataatgtaatttttatag	12093
OY	783	gaacacaaatattatcaagtataattatggttttccaataaaagaataaaaaataa	842
Db	12094	ttatttttatatgtaattttttatcatgtaattttttatgtaataataatgtaatttt	12153
OY	843	aaataatagtaagttaacaaaaaaaataattattttttccagcgccnaaaaacata	902
Db	12154	atagttatatataatgtaataataataatgtaataataataatgtaataataatgta	12213
OY	903	aaacctaaacctaaatttaaactttttagtgaaccccaaaccttggataaactttaa	962
Db	12214	atatatagttataataataatagttatataataatgta-----tatataataatgta	12269
OY	963	acattaaacattaaacactaaaccttaaaccttaaaccttaaaccttaagtgtttaa	1022
Db	12270	ataataatgtaataataataatgtaataataataatgtaataataataatgtaata	12329

OY	1023	tgattgaggttttgatttgaatttgaatttgaatttccaaagggtttaaaggtttccaaaga	1082
Db	12330	atagttatataatataatagtataatataatgtaataatataatagtatatataat	12389
OY	1083	gtctaggtttagggattatgaacttaggattagtggttttaactgacgaagttcaaagt	1142
Db	12390	atgttatgataatagtataatataatattttag---tatatgttatataata	12446
OY	1143	tttttaaaaaatttttttttgtaaacactactattttttatttttttttaacctttta	1202
Db	12447	ttttatgataattttatataataataatttbaatgataataatttataataataatttt	12506
OY	1203	tattaaaaacataaataattttaaaccacactgtttccaataagtcagtcca	1262
Db	12507	atgtatataatttatataataataattttatgataatattttatataataataattttagt	12566
OY	1263	ttatttttttgtaacaaaaaatgtgcacttagaattccaatgcgaataattatttatt	1322
Db	12567	ataattttatatacatataattttttagataatgttatataatataatataataat	12626
OY	1332	ttcagctaaataatttgcaagtgcatgtgaactctataaataattttatcccaat	1382
Db	12627	ttcgtatataataattttatataataataataataatttttcgtatataataattttata	12686
OY	1383	gctaatatgggtccaacatgtgtcaattaaatagaact-aaataatttcattttttt	1441
Db	12687	tatatataataataataatttcgtatataataattttatataataataataataatttc	12746
OY	1442	cttaactctgtgtaaaaatgtccaagttaaaatttattgaaagcaattgagtaatttt	1501
Db	12747	gataataataatttatataataataataataataatttcgtatataataatttcgtata	12806
OY	1502	gtttcattttttaaagaatatcgaatatgaataaataacacaatttttattgataagaact	1561
Db	12807	aatatataataataataataataatttcgtatataataatttcgtatataataataata	12866
OY	1562	aaaaatccccaaggaaggtgaacgcaagaataagtcacaagtttggggaagcaact	1621
Db	12867	tattattcgtatataataataatttcgtatataataataataataataatttcgtataataatt	12926
OY	1622	atggcccaaatgcatcaaaaactcttcctgtatttcaaaaactctacaatttggttag	1681
Db	12927	tcgtatataataataataataatttttataataata-----ttatataataattta	12979
OY	1682	agttaatagaccaacacacatgatactcatattagaaataattccaanaaatctactagcg	1741
Db	12980	tatatataattgtatataattttatataatgataatgtatataataataattttataatg	13039
OY	1742	aaataataaaacttctcttatttataccaatccctataaaaacttattatataact	1801
Db	13040	tattattgataataatatttttataataataattttataataattttatataataatt	13099
OY	1802	aaacaatttttaataaagaanaataagggccagatcgatcactaaanaatatatgtatt	1861
Db	13100	ttatatataataattttatataataataataatttgataataataattttatataataatttttag	13159
OY	1862	cttaagaatagatataataataataacaggtcccatattattccaataaataaataat	1921
Db	13160	ataataattttatataataataatttttagataataataattttatataataataatttttagataat	13219
OY	1922	tttcgtatgccgaataatcattactataaaattcatalaaacacacatgtagatgtattt	1981
Db	13220	atttataataataat-acttttagataataattttatataataataataatttttagataataattt	13278
OY	1982	atttatataataat-1995	
Db	13279	ataataataataatt-13292	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2002, 09:02:55 ; Search time 2155.85 Seconds
(without alignments)
17861.545 Million cell updates/sec

Title: US-09-509-188-3

Perfect score: 2853
Sequence: 1 ggaaccacacaaagaaacgcg.....ccgtcttgatttaactt 2853

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estlba:*
2: em_estlhum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estlro:*
8: em_hlc:*
9: gb_estl1:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	688.2	24.1	758	12 BH540198	BH540198 BOGTRB58TF
C 2	646.6	22.7	748	12 BH540204	BH540204 BOGTRB58TF
C 3	445.8	15.6	783	12 BH460089	BH460089 BOGSK21TF
C 4	378.2	13.3	801	12 BH563990	BH563990 BOGVZ70TR
C 5	229	8.0	872	12 BH478681	BH478681 BOGTR36TF
C 6	228.6	8.0	554	12 BH582777	BH582777 BOHNAH49TR
C 7	228.2	8.0	739	12 BH467584	BH467584 BOHRCG31TF
C 8	227	8.0	739	12 BH467584	BH467584 BOHRCG31TF
C 9	226.4	7.9	614	12 BH581638	BH581638 BOHNAH49TR
C 10	217.8	7.6	790	12 BH603601	BH603601 BOHNAH49TR
C 11	206.8	7.2	836	12 BH570528	BH570528 BOGTRB58TF
C 12	205.4	7.2	810	12 BH477804	BH477804 BOHNAH49TR
C 13	205	7.2	810	12 BH441036	BH441036 BOHRCR23TR
C 14	202.2	7.1	872	12 BH478681	BH478681 BOGTR36TF
C 15	199.6	7.0	430	12 BH544194	BH544194 BOGDD51TR
C 16	199.2	7.0	790	12 BH603601	BH603601 BOHNAH49TR
C 17	192.4	6.7	836	12 BH570528	BH570528 BOGTRB58TF

C 18	192.2	6.7	677	12 BH007803	BH007803 ee76d02.x
C 19	192.2	6.7	677	12 BH007804	BH007804 ee76d02.x
C 20	189.4	6.6	462	12 BH011048	BH011048 ep76a08.b
C 21	189.2	6.6	813	12 BH250621	BH250621 BOGAV60TR
C 22	186.4	6.5	504	9 AV549076	AV549076 AV549076
C 23	184.2	6.5	626	12 BH558350	BH558350 BOGRM31TF
C 24	182	6.4	761	12 BH425956	BH425956 BOGAV76TF
C 25	180.2	6.3	462	12 BH011048	BH011048 ep76a08.b
C 26	179	6.3	787	12 BH582306	BH582306 BOHRCR23TR
C 27	177	6.2	604	12 BH564610	BH564610 BOHRCR23TR
C 28	177	6.2	765	12 BH479185	BH479185 BOGTRB58TF
C 29	176.6	6.2	605	12 BH560885	BH560885 BOGTRB58TF
C 30	175.8	6.2	753	12 BH476224	BH476224 BOGVP19TR
C 31	172.6	6.0	703	12 BH581818	BH581818 BOGEE94TF
C 32	172.6	6.0	808	12 BH581818	BH581818 BOGEE94TF
C 33	172	6.0	812	12 BH455873	BH455873 BOHCK07TR
C 34	171.8	6.0	761	12 BH425956	BH425956 BOGAV76TF
C 35	171.6	6.0	727	12 BH570602	BH570602 BOGME46TF
C 36	171.4	6.0	785	12 BH532651	BH532651 BOHUE58TR
C 37	167.4	5.9	748	12 BH540204	BH540204 BOGTRB58TF
C 38	167	5.9	686	12 BH454053	BH454053 BOGLU21TF
C 39	166.8	5.8	796	12 BH606626	BH606626 BOGSD52TF
C 40	163.6	5.7	721	12 BH556962	BH556962 BOHLC94TF
C 41	163	5.7	614	12 BH581638	BH581638 BOHNAH49TR
C 42	162.4	5.7	799	12 BH421537	BH421537 BOGIG23TR
C 43	161.8	5.7	746	12 BH592365	BH592365 BOHMA61TF
C 44	160.8	5.6	576	12 BH456908	BH456908 BOGUX51TR
C 45	160.8	5.6	603	12 BH010155	BH010155 ep57g11.b

ALIGNMENTS

RESULT 1
BH540198/c
LOCUS BH540198 758 bp DNA linear GSS 14-DEC-2001
DEFINITION BOGTRB58TF BOGT Brassica oleracea genomic clone BOGTRB58, DNA sequence.
ACCESSION BH540198
VERSION BH540198.1 GI:17785802
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 758)
TOWN,C.D., Van Aken,S., Uterback,T. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BOGTRB58TR
COMMENT
Contact: Chris Town

REFERENCE
TOWN,C.D., Van Aken,S., Uterback,T. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BOGTRB58TR
COMMENT
Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: ctown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
Location/Qualifiers

FEATURES
source
1..758
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone_lib="BOGT"
/note="Vector: pGEM1, site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pGEM1 using BstXI linkers"

BASE COUNT
228 a 106 c 155 g 269 t

|||||
Db 671 CAAATTTTATTTTATTTTACCTTTTATAT-T-AAACATATATATTTTACTCCA 729
Oy 1234 tctgttcatltaa 1248
|||||
Db 730 TCTGTTTCATATTA 744

RESULT 3
BH460089 783 bp DNA linear GSS 13-DEC-2001
LOCUS BOGSK21TF BOG5 Brassica oleracea genomic clone BOGSK21, DNA
DEFINITION sequence.
ACCESSION BH460089
VERSION BH460089.1 GI:17649834
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 783)
Town, C.D., Van Aken, S., Uterback, T. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BOGSK21TF
COMMENT Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.

FEATURES
Location/Qualifiers
source 1..783
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGSK21"
/clone_1lb="BOG5"
/note="Vector: PHOS1, Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 228 a 183 c 161 g 211 t
ORIGIN

Query Match 15.6%; Score 445.8; DB 12; Length 783;
Best Local Similarity 92.3%; Pred. No. 2.4e-38;
Matches 481; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

Oy 2101 ttgttccttgatgaccccttctaagctggtgacatcccaagttaagcaagaac 2160
|||||
Db 1 ttgttcttctatgaccccttcttcaatgagtgctacatccaaagcagcaac 60

Oy 2161 tagagtcacaaactaaacaaagacgtcttctatcgcgcactactgcttaccacca 2220
Db 61 TAGAGTCATCACTAACCAAGTGTCTTCTATCGGGGCACTTGCCTTCACCCCA 120

Oy 2221 agcaacattgagcgttcggtggtcgcgaagaagccttccctgagggcacttcgactaa 2280
Db 121 AGCCACATTTGGAGTTCCTGCGCTCGGAAAGCCCTTCCTCGAGGCCACTTCGACAA 180

Oy 2281 ctcctgtcatctgcacagaagatgacacagatgcttgaacgaagaaggagtgagta 2340
Db 181 CTCCTGTCATCTGCGCACAGGAAGTCAACAGATGCTTGTCGCCAAGAAGAGGTAGTA 240

Oy 2341 catgttttaagatagatcgctgaagcttcttccacagaagaagcgcgtattgagtcgaat 2400
Db 241 CAGTGTGATGATATCGTTGAGACTTTCTCTCCACAGGAAGCGCGCTATTGAGTGGAAAT 300

Oy 2401 gttagcgcgatcacaagaagatgagcaagaagtgtgagaagaccgtcttgatcttcc 2460

|||||
Db 301 GTTGCCCGCCGATCAAGAGATGAACAAAGATTGTGAGAAGACCGCTTTGGATCTTCC 360
Oy 2461 atgaccctcttgaccgctatgtaagctatattgctccaccgttgttgactact 2520
Db 361 ATGACCCCTTCTTGACAGGCTATGTCAAACTATATGCTCCACCGGTGTGGATCTACT 420

Oy 2521 caectctcttccacagctccttataatgctctcttccacagagctccttccatgctc 2580
Db 421 CACCTCTCTCTTACATGCTCTTCACAGGCTC---CTTACATGCTCTTCACAGGCTC 477

Oy 2581 ctccacatgctccttccacagagctccttataatgctccttca 2621
Db 478 CTTTACTGCCCGCTTCACAGCTCCTCCACCGGCTCAGTGA 518

RESULT 4
BH563990 801 bp DNA linear GSS 14-DEC-2001
LOCUS BOGVZ70TF BOGV Brassica oleracea genomic clone BOGVZ70, DNA
DEFINITION sequence.
ACCESSION BH563990
VERSION BH563990.1 GI:17815830
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 801)
Town, C.D., Van Aken, S., Uterback, T. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BOGVZ70TF
COMMENT Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.

FEATURES
Location/Qualifiers
source 1..801
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGVZ70"
/clone_1lb="BOGV"
/note="Vector: PHOS1, Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 231 a 176 c 177 g 217 t
ORIGIN

Query Match 13.3%; Score 378.2; DB 12; Length 801;
Best Local Similarity 99.2%; Pred. No. 3e-31;
Matches 380; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2114 gatcctcttcttaagctggttccatccaaagttaagcaagaactagatcatcaac 2173
Db 1 GATCCATTCTTTAAGCTGTGTACATCAACAAAGTTACAGCAAGAACTAGATCATCAAC 60

Oy 2174 taaccaagagcttcttctatcgcgcactactgcttaccaccaagaagaattggcc 2233
Db 61 TAACCAAGAGCTCTTCTTATCGGCACTTACCTGCTTCACCCCAAGCAACATTTGGCC 120

Oy 2234 gtccgtgctcgcgaagaagccttccctgagggcacttcgactaactcgttccatc 2293
Db 121 GTTCGTGCTGCTCGGAAGACCTTCCTCGAGGTCACTTCGCACTAATCTCCATCATC 180

Oy 2294 gccacaggaatcaccaagatgcttgaacgaagaaggagtagtatacgtttaaagta 2353

Db 181 GCCACGAGAGTCACCAATCTTGAAGCAGACAAAGAGTACGTCATGTTTAATGA 240
OY 2354 tatcgctgaagctttctcaccaggaagccgctatttgatcggaaagtgtgcgcgcgat 2413
Db 241 TATGCTTACAGCTTTCTTCCACCGAAAGCCGCTATTGATCGGAATGTTGCCCGGAT 300
OY 2414 caagaagatgacaaagatgtggaagacgctcttgatctttcattgaccccttct 2473
Db 301 CAAAGAGATGAAACAAAGATGTGAGAGACCGCTTTGATCTTTCAATGACCCCTTCTT 360
OY 2474 gaccggtatgtcaagctacatt 2496
Db 361 GACCGCTATGTCAAGCTACATT 383
RESULT 5
BH478681
LOCUS BH478681 872 bp DNA linear GSS 13-DEC-2001
DEFINITION BOGT236TF BOGT Brassica oleracea genomic clone BOGT236, DNA
sequence.
ACCESSION BH478681
VERSION BH478681.1 GI:17686785
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE Town, C.D., Van Aken, S., Utecher, T. and Fraser, C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BOGT236TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
FEATURES
source location/Qualifiers
1..872
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGT236"
/note="Vector: PHOS1, Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 313 a 135 c 109 g 315 t
ORIGIN
Query Match 8.0%; Score 229; DB 12; Length 872;
Best Local Similarity 72.2%; Pred. No. 1.3e-15;
Matches 390; Conservative 0; Mismatches 121; Indels 29; Gaps 6;
OY 735 ccaacataaattcattatgcatatcattatcatttgaagatgaacaaat 794
Db 4 CCAACCAATGATTCATCTACTCTATATTCATATTTTGTAGAAAGTTTACAAATAT 63
OY 795 tatcaagtatattatgttttccaataaaagataaaataaaataaaataaata 854
Db 64 TGTCAAGTATATATTAGCTTTTAAATTAAGATTAATAAATAAATAAATAAATA 123
OY 855 gttaacaaaaaataaataaataaataaataaataaataaataaataaataaata 910
Db 124 GTTACAAAATAATATATTAATAAATAATTCAGTCTGTAGCAAAACACTTAACCTTAA 183
OY 911 accctaataaataaataaataaataaataaataaataaataaataaataaata 965

Db 184 CTCCTGATATACCTCTTAACCTTTGGTAAATCTTAACCTTTGATTAATCTTAACCTC 243
OY 966 ttaaactaaataaataaataaataaataaataaataaataaataaataaataaata 1025
Db 244 TAAATCA-AAAACACTTAACCTTAACCCMAATCTTGAATATTTTGTGTTAAATGTT 301
OY 1026 ttaagtgtt-----ttgattatagtttagattatccaaggttaagttt 1074
Db 302 TTGATTTAGAGTTTACGATTTTATCCAAAGGTTTAGGCTTACCTTAAGGTTAGGCTTT 361
OY 1075 acccaagattatggttaggattatgacttaggattataggttttctagacgctt 1134
Db 362 ATCCAGGTTTAAAGATATGATTTAGAGTTTACGTTTACGTTTACGTTTACGTTTACG 421
OY 1135 caaagtattttaaataaataaataaataaataaataaataaataaataaataaata 1188
Db 422 AAAATATGCTTTTAAATAATTTCTAATTTTGTGTTTACCTATTTCTATTTTATTTT 481
OY 1189 tttaacttttataataaataaataaataaataaataaataaataaataaataaata 1248
Db 482 TTTTAA-TTTTGTGTTTAAATAAATAATTAATTTGATATATTTGTTTTTAAATTA 540
RESULT 6
BH582777
LOCUS BH582777 554 bp DNA linear GSS 15-DEC-2001
DEFINITION BOHNA49TR BOHA Brassica oleracea genomic clone BOHNA49, DNA
sequence.
ACCESSION BH582777
VERSION BH582777.1 GI:17835234
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE Town, C.D., Van Aken, S., Utecher, T. and Fraser, C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BOHNA49TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
source location/Qualifiers
1..554
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOHNA49"
/note="Vector: PHOS1, Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 204 a 88 c 70 g 192 t
ORIGIN
Query Match 8.0%; Score 228.6; DB 12; Length 554;
Best Local Similarity 73.9%; Pred. No. 1.8e-15;
Matches 343; Conservative 0; Mismatches 90; Indels 31; Gaps 3;
OY 718 cttaaccttaagtttcccaacataaattcattatgcatatcattatcattttag 777
Db 69 CTTCACCTCTAGGTTCCACCAACCAATAGATTTTCCTTATTTCAAAATTCGGATATCTTTGG 128
OY 778 aaagtgaacaaataaataaataaataaataaataaataaataaataaataaata 837

BASE COUNT 263 a 95 c 99 g 282 t

Query Match	8.0%;	Score 228.2;	DB 12;	Length 739;
Best Local Similarity	68.8%;	Pred. No. 1.7e-15;		
Matches 407;	Conservative 0;	Mismatches 139;	Indels 46;	Gaps

FEATURES

SOURCE

1..739

FEATURES	Location/Qualifiers
source	1..739
	/organism="Brassica oleracea"
	/strain="TO1000DH3"
	/db_xref="taxon:3712"
	/clone="BORHG31"

/clone.lib="BOHR"
 /note="Vector: PHOS1; Site:1: BstXI; 2-3 kb sheared
 genomic DNA inserted into PHOS1 using BstXI linkers"
 BASE COUNT 263 a 95 c 99 g 282 t
 ORIGIN

Query Match 8.0%; Score 227; DB 12; Length 739;
 Best Local Similarity 73.0%; Pred. No. 2.3e-15;
 Matches 373; Conservative 0; Mismatches 105; Indels 33; Gaps 5;

Y 725 tctaatgctcaccacaaataaattcattatgcatattctatctttagaagttg 784
 DB 685 TCTAGATTCAACCACTAATAGATTATATTTCAATTCGATCTTTAAATGGA 626
 Y 785 aacaaatattcaagttattatgitttttcaataaagaataaataaataa 844
 DB 625 AACAAATATTAT-ACCTTATTATTGTTTAAATAAGCAAGTAAACATAAATAA 567
 Y 845 aataatagtagttacaaaaaataaataatttttaccagcgctcnaaaaccta 904
 DB 566 AATAAGAGGAGTTACATATAATTAATTTTA-----AAAAATAT 526
 Y 905 acctaaaccttaataatttaacatttagttaaaccttaaaccttggataa 964
 DB 525 TTAAAGCTGCTCAATAAATTAATTCCTAAACATTAAGCCCTTGAGTAAATCCTAAC 466
 Y 965 attaacattaaacacctaaccctaaccctaaccctaaccctaacccta 1024
 DB 465 TCTAATCAAAATACTAATAACCTAATAACCTTAAGCTGTTTAAAGT 406
 Y 1025 tttaagttttttagatttaagtttagatttaccgaagtttaagtttaagtt 1084
 DB 405 TTTAGTGTGTTTGAATTAACCTTAGATTATCCAGGTTTATGTTTACC----- 351
 Y 1085 ttatggttaggaattatgaattagtttagttttttaggaagtttaagtt 1143
 DB 352 -AAGGTTTAGGATTAGGATTAGATTAGTTGTTTGGTGAAGCGCTAAATAATTA 294
 Y 1144 tttaaaaaatt 1199
 DB 293 TTTAAAAAACATTTTGTGTAACCTACTACTATTTTATTATTATTATTATTATCTTT 234
 Y 1200 ttatatataaacataataataattatct 1230
 DB 233 TTATTTTAAAAACATATAATTAATCTGAATAAT 203

RESULT 9
 BH581638/c 614 bp DNA linear GSS 15-DEC-2001
 LOCUS BOHNM45TR BOHL Brassica oleracea genomic clone BOHNM45, DNA
 DEFINITION sequence.
 ACCESSION BH581638
 VERSION BH581638.1 GI:17833919
 KEYWORDS GSS.
 SOURCE Brassica oleracea.
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 614)
 Town, C.D., Van Aken, S., Uiterback, T. and Fraser, C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 unpublished (2001)
 Other GSSs: BOHNM45TR
 COMMENT Contact: Chris Town
 TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF
 Class: sheared ends.
 Location/Qualifiers
 source 1..614
 /organism="Brassica oleracea"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BOHNM45"
 /note="Vector: PHOS1; Site:1: BstXI; 2-3 kb sheared
 genomic DNA inserted into PHOS1 using BstXI linkers"
 BASE COUNT 225 a 80 c 89 g 220 t
 ORIGIN

Query Match 7.9%; Score 226.4; DB 12; Length 614;
 Best Local Similarity 71.0%; Pred. No. 2.9e-15;
 Matches 387; Conservative 0; Mismatches 117; Indels 41; Gaps 5;

Y 719 taaccttaagttcaccacaaataaattcattatgcatattctatctttaga 778
 DB 588 TGAACCTCTAGTCAACCAACCTATAGATTATTAATTAATTAATTAATTA 529
 Y 779 aagtgaacaaatattatcaagttatattatgatttttcaataaagaataa 838
 DB 528 ATAGGAACCAAAATATGTCAGATTATTAATTAATTAATTAATTAATTA 481
 Y 839 ataaaaataatagtttagttacaaaaaataaataatttttaccagcgctcna 898
 DB 480 GTAAACCAAAATATGTTATTAACGAAATAAATAATTAATTAATTAATTA 430
 Y 899 actaaacctaaccctaataataatttaacatttagttaaaccttaacatt 958
 DB 429 ACTAAACCTTAATCCTATTCCTAACTCTGGGTAACCTTAACCTTAATTAAT 370
 Y 959 ttaaa-----cattaaacttaaaacctaaccctaaccctaaccct 1003
 DB 369 CTAAACACTAAATTAATAAAGCACTAAGCACTAAACCTTAATTAATTAAT 310
 Y 1004 aaaccttaagttttaaatttagtttttttttttttttttttttttttttt 1063
 DB 309 AAACCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 250
 Y 1064 gtttaagtttaccacaaagtttagtttagtttagtttagtttagtttagtt 1123
 DB 249 GTTTAGGTTTATCCAGGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 190
 Y 1124 ctgacgacgttcaaaagtttttttaaaaaataatttttttttttttttt 1183
 DB 189 GTTTAG-TGTTTGTGCGCAGCAGCTTAATAAATAATTTTGTATTAATTA 135
 Y 1184 ttatt 1243
 DB 134 ATTCTTTTACCTTTTAATTTAATAAATAATTAATTAATTAATTAATTA 75
 Y 1244 attaa 1248
 DB 74 TTAA 70

RESULT 10
 BH603601/c 790 bp DNA linear GSS 15-DEC-2001
 LOCUS BOHLM68TR BOHL Brassica oleracea genomic clone BOHLM68, DNA
 DEFINITION sequence.
 ACCESSION BH603601
 VERSION BH603601.1 GI:17856047
 KEYWORDS GSS.
 SOURCE Brassica oleracea.
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.

RESULT	12
BH477804/c	
LOCUS	BH477804
DEFINITION	BH477804 810 bp DNA
	linear GSS 13-DEC-2001
	BOHNS55cTR BOHA Brassica oleracea genomic clone BOHNS55, DNA sequence.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 810)	Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.	Whole genome shotgun sequencing of <i>Brassica oleracea</i>	Unpublished (2001)	Other_GSSs: BOHA5567F

Contact: Chris Town
TIGR
9712 Medical Center Drive,
Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES	
SOURCE	Location/Qualifiers
	1..810
	/organism="Brassica oleracea"
	/strain="T01000DH3"
	/db_xref="taxon:3712"
	/clone="BOHAS36"
	/clone_1ib="BOHA"
	/note="Vector: pHOS1, site_1: BstXI, 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT	280 a 95 c 123 g 312 t
ORIGIN	

Query Match	7.28;	Score 205.4;	DB 12;	Length 810;
Best Local Similarity	67.7%;	Pred. No. 4.1e-13;		
Matches 410;	Conservative	0;	Mismatches 127;	Indels 69;
			Gaps	6

607 AACCTCTAAATTTCACCAAGCCCATTAATATTTTCTATTGATGTGGCATAATCTATATCTTTGAAGA 780
AACCTCTAAGTTCACCACAACAATAAATCTCATTTATGTGCATATCTTATGAAA

[illegible]

547 AAGAAACCAATATATTCAAATATATTAAGTTTAAATATAAAAACTGAAAAAATAGC 488

b
y
650 -----aaataatataatatagtagttacaaaaaaaattataattt 880
||||| ||||| ||| | | | |
487 AGCTACAGAAAAAAGATTAAATAAAAATTTTAACTTCGCGTAATCAATCCTT

881 ttaccagcgtcanaanaaactaaactatgaacc-----taataatftaacit 928
 428tgcacatgcgttcgcaaacacacacitaaacccct 428

427 AATCCCTAATCCCTAAACCCCTAATCTTAACCCCTAAACCCCTGGGTAAACCATTAACCC 368

```

929 ctagttaaccctaaccttgataatcgt-----aaacattacaactta 975
||||| ||||| ||||| |
||||| ||||| ||||| |
||||| ||||| ||||| |
||||| ||||| ||||| |

```

QY 976 aaacactaaaccctaatactcctaactcctaacccttaagtgtttaatgtttagtgtttt 103

Db 307 AACCTAAACCTAAATCCTAAACCCCTAAACCCCTTGAGTGTTTAGTGTATTGATTT 248

Db 247 TGATTAGAGTTAAGATTATCTAGAGTTAAGATTATCTAGAGTTAAGAGTTAGGGTTAC 188

QY 1095 -----gggattatgacttagatttagtcttactgcagcagcttcaagtaatt 114

Db 187 CCAAGTGTTCGGCGTTAGCAATTAGCGTTTTCAGGATTTAGGATTTAGGCTTAATGTTT 128

[illegible]

QY .1192 ttaccttttataataaacataataataatttaatactccatctgtttcattatgaagtg .125

DB	67	TTACCTTTTAATTTTAAAGATATATATTAATTGACCAATATTTGTTAACTTTTAAAG	8
QY	1252	tcattg	1257

Db 7 ATATCG 2

RESULT 13

LOCUS	810 bp	DNA	linear	GSS 12-DEC-
BH441036				
BH441036				
BOHCR3TR	ROHC	Brassica oleracea		
DEFINITION				

[illegible]

KEYWORDS	BRASSICA OLERACEA
SOURCE	GSI:1/b26/50
BRASSICA OLERACEA.	

ORGANISM
Brassicaceae
Eukaryota: Viridiplantae: Streptophyta; Embryophyta: Tracheophyta
Spermatophyta: Magnoliophyta: eudicotyledons

REFERENCE
Rosidae; eunotsids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 810)

TITLE Whole genome shotgun sequencing of *Brassica oleracea*

COMMENT
Contact: Chris Town
TIGR
9712 Medical Center Drive Rockville MD 20850 nra

Tel: 301-838-3523
Fax: 301-838-0208
Email: cd@cmu.edu

URL: eucoweb119.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR

FEATURES	Class: sheared ends.
source	location/Qualifiers
	1. .810

```

/organism="Brassica oleracea"
/strain="T01000DH3"
/db_vref="#taxon.3710"

```

```

/case_art="BOHCRC23"
/clone_lib="BOHC"

```

```

/note=Vector: PHOS1: Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"
260 a 200 c 141 g 209 t

```

ORIGIN

Query Match	7.2%	Score 205;	DB 12;	Length 810;
Best Local Similarity	93.0%	Pred. No. 4.5e-13;		

Matches	226;	Conservative	0;	Mismatches	15;	Indels	2;	Gaps
---------	------	--------------	----	------------	-----	--------	----	------

50 GGGTCCACAGAGAAACCGAGAGAGTAAAGTTCCGACCTTCATTAATATATATTTGT 109

QY	61	ttcagccatcatcatgatacaaacactcccttagccatcatgataaagtgtgtttag	120
Db	110	ttcacc--cattcaatgtaacaaagaaacttttagcttcacagggatataaagtgtgtttag	167
QY	121	attcaataaccgggttcgagtcataagacttgcaacttttcacactttttaaagttag	180
Db	168	atcccaataaacccgggttcgagtcacacagacttgacactttttcacattttttaaagttag	227
QY	181	aacgcacatcatgcgtcagctgtcgcataagagatgacgaactgtctataataatgtag	240
Db	228	aaccacacatcatgcgtcagcgtgcgcataagagatgacgaactgtctataataatgtag	287
QY	241	att 243	
Db	288	att 290	
RESULT 14			
LOCUS	BH478681/c	872 bp	DNA linear GSS 13-DEC-2001
DEFINITION	BH478681	BOGT236TF BOGT Brassica oleracea genomic clone BOGT236, DNA	
ACCESSION	BH478681		
VERSION	BH478681.1	GI:17686785	
KEYWORDS	GSS.		
SOURCE	Brassica oleracea.		
ORGANISM	Brassica oleracea.		
REFERENCE	Enatryola; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.		
AUTHORS	1 (bases 1 to 872)		
TITLE	Town, C.D., Van Aken, S., Uterback, T. and Fraser, C.M.		
JOURNAL	Whole genome shotgun sequencing of Brassica oleracea unpublished (2001)		
COMMENT	Other GSSs: BOGT236TR		
	Contact: Chris Town		
	TIGR		
	9712 Medical Center Drive, Rockville, MD 20850, USA.		
	Tel: 301-838-3523		
	Fax: 301-838-0208		
	Email: cdtown@tigr.org		
	DNA is from a doubled haploid provided by Tom Osborn.		
	Seq primer: TF		
	Class: sheared ends.		
FEATURES	Location/Qualifiers		
SOURCE	1..872		
	/organism="Brassica oleracea"		
	/strain="TO1000DH3"		
	/db_xref="taxon:37112"		
	/clone="BOGT236"		
	/clone.lib="BOGT"		
	/note="vector: pHD1, site 1: BstXI, 2-3 kb sheared genomic DNA inserted into pHD1 using BstXI linkers"		
BASE COUNT	313 a 135 c 109 g 315 t		
ORIGIN			
Query Match	7.1%	Score 202.2:	DB 12; Length 872;
Best Local Similarity	70.9%	Freq. No. 8.5e-13;	
Matches 380; Conservative	0;	Mismatches 119;	Indels 37; Gaps
QY	725	ctcaagttaccacaacaataattcatattgatcatatcattcatatcattttagaagtga	784
Db	590	tctatatttcacacacactcactaagactctgtgtatttcattatcattttaaataaaaa	531
QY	785	aacaaatatttcaagttatattatgttttcaaataaagaagataaaaaataataaa	844
Db	530	aacaaatatttcaagttatattatgtttttaaataaaaaa-ataaaaaaaataaaaa	472
QY	845	aataatagtaagtcaaaaaaataataataattttccagcgt-----	890
Db	471	aatagtaagtcaaaaaaataataataatttttttaaagcctatttttaacgtcgtca	412

OY	891	-----canaaacactaaaccttaacccttaataattatnaacttttagtlaaaccttaa	944
DB	411	GCAAAACCTTAACCCCTTAAACTCATATCAATCAATCCTTTAAACCCTTGATTAACCCCTTAA	352
OY	945	cctttgatatcaactttaacactt-----aaacattaacaacccaacccctaactcctaac	1000
DB	351	COCCTTAGGTAACCCCTTAAACCTTTGGATTAAATCCTAACCTCAATAATCAAAAACATTAAAC	292
OY	1001	tctaaaccttaagtgttta-----aatgttagtgltttgattagtttagttgatt	1054
DB	291	AAAAAATACCTCAAGGATTATAGGTTCAGTTTACGTTTACGTTTTCATTGATTGAAG--TTAGGATT	233
OY	1055	tatccaaaggatttaaggtttaccagaagatttaagtttagtgattatgaacttagatt	1114
DB	232	TATCCAAAGGTTTAGATTATACCAAGAATTAGAGCT--GTAATCAACAGTTTAGGTTT	175
OY	1115	agtgtttactgacagcgftcaaglatcttttaaaaaatatlttttgtlaacaactac	1174
DB	174	AOTGTTTTGCCTAACGAC--CTGAATTTTTTTTATATATATTTTTTTTGTAACTAGTAC	118
OY	1175	tattttattatcttttttaccttttatataaaaacaataataatlaact	1230
DB	117	TATTTTTATTTATTTATTTTATCTTTTATTTTAAACCTAATATATACCTTGACAAT	62

RESULT 15
BH544194

```

LOCUS      BH544194
DEFINITION BOGDD51TR BOGD Brassica oleracea genomic clone BOGDD51, DNA
sequence.
ACCESSION  BH544194
VERSION    BH544194.1 GI:17795975
KEYWORDS   GSS.
SOURCE     Brassica oleracea.
            Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1 (bases 1 to 430)
AUTHORS     Town,C.D., Van Aken,S., Uterback,T. and Fraser,C.M.
TITLE       Whole genome shotgun sequencing of Brassica oleracea
            unpublished (2001)
JOURNAL     Other GSSS: BOGDD51TF
COMMENT     Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
```

FEATURES

source Location/Qualifiers

1..430

/organism="Brassica oleracea"

/strain="TO100DH3"

/db_xref="taxon:3712"

/clone="BOGDD51"

/clone.lib="BOGD"

/note="Vector: PHOS1, site.1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 168 a 66 c 46 g 150 t

ORIGIN

Query Match 7.0%; Score 199.6; DB 12; Length 430;

Best Local Similarity 76.9%; Pred. No. 2.3e-12;

Matches 270; Conservative 0; Mismatches 75; Indels 6; Gaps 2.

OY	719	taaacctcaagttcaccaacaataaaattcatatgatcatatctatctttaga	778
DB	80	TCATACCTTAGGTAGTACCAACCAACATGATCGTATTTTATTTTGTATTCCTTTTACA	139

```

QY 779 aagtgaaacaaatattatcaggtatattatgttttcaataaagaataaataa 838
Db 140 AAAGAAACAAATATGTCAAGTTATATGTTTTTAAATPAAAGGTAAATAA 199
QY 839 ataaaaataatagtagttcaaaaaaaataatatttaccagctc---ana 894
Db 200 AATPAAATATATATAGTTACAAATTTTTTTTTTAAATATATTTTTTAGGTCGTCATCA 259
QY 895 aaacactaaacccaacccaataatttaacttttagtgaacccaacccaaccca 952
Db 260 AAACACTAAACCCCTAAATCCTAAACCAAGAAATCTTTGGTAAACCCCTAAATCCTAGCTC 319
QY 953 taaacttaaacattaaacattaaacccaacccaacccaacccaacccaaccca 1012
Db 320 TAAATPAAACATTAACATTAACATTAACATTAACATTAACATTAACATTAAC 379
QY 1013 agtgtttaagtgtgtgtttgtatcatagtttaggattatccaag 1063
Db 380 ACCCTTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 430

```

Search completed: June 28, 2002, 12:25:26
 Job time: 12151 sec